

K. RIBNIKOV

ANALISIS COMBINATORIO

Análisis combinatorio

К.А.Рыбников

**ВВЕДЕНИЕ
В КОМБИНАТОРНЫЙ АНАЛИЗ**

Издательство Московского университета

K. RÍBNIKOV

Análisis combinatorio



Editorial Mir Moscú

Traducido del ruso por K.Medkov

Impreso en la URSS

На испанском языке

ISBN 5-03-000610-9

© Издательство Московского университета, 1985
© traducido al español, editorial Mir, 1988

ÍNDICE

Prólogo	6
Capítulo 1. Fundamentos teóricos del análisis combinatorio	7
1.1. Qué se estudia en el análisis combinatorio y qué clase de problemas se resuelven	7
1.2. Conocimientos indispensables de la teoría de los conjuntos y del álgebra	9
1.3. Muestras y ordenaciones	17
1.4. Distribuciones y llenados	23
1.5. Sistemas de conjuntos	32
Capítulo 2. Funciones generatrices	37
2.1. Fundamentos del método de funciones generatrices	37
2.2. Tipos de funciones generatrices y de numeradores	40
2.3. Aparato operacional del método de funciones generatrices	52
2.4. Sobre las aplicaciones del método de funciones generatrices	62
2.5. Teoría de Redfield-Polya	66
Capítulo 3. Métodos lógicos	74
3.1. Método de inclusiones y exclusiones	74
3.2. Sistemas de representantes de los conjuntos	78
3.3. Principios de la teoría de Ramsey	83
Capítulo 4. Aparato tabular de matriz del análisis combinatorio	87
4.1. Sistemas de incidencia y matrices especiales	87
4.2. Rectángulos y cuadrados latinos	93
4.3. Bloque-esquemas	101
4.4. Permanentes	113
Capítulo 5. Medios geométricos y gráficos	130
5.1. Evidencia en las interpretaciones de los sistemas discretos	130
5.2. Ideas geométricas finitas	132
5.3. Sobre los grafos	144
Capítulo 6. Métodos de resolución de los problemas extremales	159
6.1. Planteamiento de los problemas combinatorios extremales y accesos a su resolución	159
6.2. Método de ramificaciones y restricciones ¹⁾	165
6.3. Métodos heurísticos	179
6.4. Optimización en los grafos	189
6.5. Flujos en las redes	192
Capítulo 7. Métodos probabilísticos en el análisis combinatorio	206
7.1. Ejemplos de aplicación de los métodos probabilísticos	206
7.2. Problemas de planificación del experimento	210
7.3. Método de entropía	214
7.4. Método de balance aleatorio	220
7.5. Sistemas separadores de subconjuntos	224
Capítulo 8. Análisis combinatorio en los conjuntos parcialmente ordenados	232
8.1. Conjuntos parcialmente ordenados	232
8.2. Retículos	263
8.3. Funciones de incidencia e inversión de Moebius	248
8.4. Matroides	324
Bibliografía	361
Índice alfabético de materias	367

PRÓLOGO

En el presente libro se examinan los métodos matemáticos de investigación de los sistemas discretos. Esta clase de sistemas ocupan un notable lugar en la actividad de la gente y existe una diversidad de tipos difícil de examinar (circuitos eléctricos, flujos de transporte e información, sistemas de organización de la producción, tablas, grafos, razonamientos lógicos, etc.).

Pese a que el carácter específico y heterogéneo de los sistemas discretos aparenta ser insuperable, éstos tienen en sus fundamentos mucho de común. Esto se manifiesta con toda evidencia en la semejanza e, incluso, en la coincidencia de los modelos matemáticos que se introducen al resolver problemas de carácter discreto.

Las amplias clases de los sistemas discretos se analizan en nuestro libro como partes de una teoría única. Dentro de los márgenes de dicha teoría (que históricamente conserva su nombre de análisis combinatorio) las clases citadas ocupan su lugar en dependencia de cómo se interpretan sus entes, qué problemas se estudian y cuáles son los métodos que se eligen para resolverlos. Tal exposición, como muestra la experiencia, facilita el estudio del análisis combinatorio y atribuye a esta parte de las matemáticas los rasgos necesarios de universalidad, tanto teórica como práctica.

Para que el lector obtenga una mejor orientación en la moderna teoría combinatoria general, el capítulo ocho del libro, escrito en colaboración con A.M.Reviakin, contiene material referente a la aplicación de problemas y métodos del análisis combinatorio a los conjuntos de naturaleza lo más amplia posible.

Resta por añadir que el autor expresa su más profundo agradecimiento a sus discípulos y colegas que prestaron una valiosa cooperación, y muy especialmente a A.M.Reviakin, quien ayudó activamente al autor en el proceso de preparación de la versión española del libro.

El autor quedará agradecido por las observaciones de los lectores.

CAPÍTULO I FUNDAMENTOS TEÓRICOS DEL ANÁLISIS COMBINATORIO

En este capítulo se dan explicaciones que tienen por fin permitir al lector la formación de nociones elementales sobre el análisis combinatorio (combinatoria). Se trata de los objetos que se estudian en el análisis combinatorio (considerado como una asignatura) y de los problemas que en este caso se plantean y se resuelven. Además, se introducen los conceptos principales, operaciones y simbolismo. Por fin, se explican los procedimientos aplicados en la resolución de los problemas combinatorios no complejos que se plantean tanto en la teoría como en la práctica.

1.1. QUÉ SE ESTUDIA EN EL ANÁLISIS COMBINATORIO Y QUÉ CLASE DE PROBLEMAS SE RESUELVEN

Enunciemos, para iniciar, unos cuantos problemas. Harémoslo para que el lector perciba intuitivamente el carácter combinatorio de dichos problemas y esté mejor preparado para poder asimilar las formulaciones generales.

a) Uno que ha terminado la enseñanza media y aspira a ingresar en la Universidad tiene que aprobar cuatro exámenes, siendo el sistema de calificación de cinco puntos. Para ingresar es suficiente obtener 17 puntos. ¿Cuántas son las combinaciones que le permiten aprobar los exámenes (por supuesto, sin sacar ni un solo mal)?

b) ¿Cómo se debe buscar el itinerario más corto para un cartero o un operador cinematográfico que atienden un número dado de poblados? La distancia entre cada par de poblados se conoce.

c) ¿Qué número de reinas o de otras piezas de ajedrez es suficiente para que ellas mantengan bajo amenaza todas las casillas del tablero de ajedrez y no se puedan comer una a la otra? ¿Cuántas son las combinaciones de su colocación?

d) ¿En cuántas partes dividen un espacio n planos, si de cuatro cualquiera de ellos ninguno pasa por un mismo punto, de tres, ninguno pasa por una misma recta y de dos, ninguno es paralelo, mientras que cualesquiera tres planos tienen un punto común?

Los problemas de esta índole son semejantes por el hecho de que, en primer lugar, en ellos se estudian los conjuntos discretos (compuestos de elementos separados, aislados). En la mayoría de los casos los conjuntos mencionados son finitos. Tampoco se excluyen del análisis combinatorio los conjuntos infinitos, compuestos por un número infinitamente grande

de elementos, siempre que haya una información suficiente sobre las peculiaridades estructurales de estos últimos conjuntos. La estructura de los conjuntos discretos puede ser muy compleja en función de las relaciones y razones existentes entre ellos. La primera principal tarea del análisis combinatorio consiste en estudiar tales estructuras discretas y expresar sus propiedades empleando los métodos adoptados en las matemáticas (analíticos, gráficos, tabulares y geométricos).

Sobre los conjuntos discretos se realizan operaciones. Algunas de ellas originan el cambio de la estructura de los conjuntos, otras modifican la composición de éstos. Como operaciones más simples del primer tipo intervienen las permutaciones habituales de elementos, y del segundo tipo, la separación de los subconjuntos de elementos, o, como suele decirse, sus muestras. Como regla, al resolver los problemas, las operaciones se aplican más de una vez (reiteradamente) y, además, en las combinaciones más diversas, cuando vienen impuestas diferentes condiciones. Esto es una razón por la que se prestan posibilidades prácticamente inagotables de crear construcciones discretas, las cuales se denominan frecuentemente configuraciones (a veces, configuraciones combinatorias).

Según sean el carácter del objeto de investigación y las operaciones a introducir, se determina la totalidad específica de los problemas que se resuelven mediante el análisis combinatorio. Los más antiguos son problemas sobre el número de construcciones discretas que satisfacen las condiciones planteadas. Los métodos de su resolución han recibido el nombre de problemas enumerativos. Además de los problemas de tipo enumerativo, en el análisis combinatorio se estudian cuestiones de existencia o no existencia de una configuración que satisfaga las condiciones prefijadas, se buscan algoritmos de construcción de las configuraciones, como también las cuestiones de selección en una totalidad dada de tales configuraciones o algoritmos que posean, en un mayor o menor grado, las propiedades elegidas (problemas y métodos de optimización).

El análisis combinatorio está asociado con muchos apartados de las matemáticas e incluso tiene partes comunes correlacionadas. Efectivamente, los elementos de los razonamientos combinatorios surgieron en tiempos remotos, antiguamente, en la aurora de la formación de las ciencias matemáticas. No obstante, en el transcurso de la historia estos elementos se desarrollaban junto con otras partes de las matemáticas, integrando en diversos casos la composición de éstas. No es difícil ver que en una serie de ramas de las matemáticas modernas (teoría de los números, álgebra, geometría, lógica matemática y otras) muchos conceptos y métodos fundamentales tienen naturaleza discreta y poseen conexiones estables. Esto permite estudiar los problemas del análisis combinatorio en diferentes interpretaciones, investigar problemas, que a primera vista parecen tener diferente naturaleza, desde un punto de vista único que corresponda del modo más conveniente a la esencia del problema.

Hoy día las posibilidades de los métodos discretos de investigación aumentaron muy bruscamente. Ha crecido también la importancia de estos métodos. A la par con la teoría combinatoria general formada históricamente (análisis combinatorio) existen en las matemáticas modernas: la teoría de los grafos y de los hipergrafos, geometría de los números, análisis discreto y análisis finito, investigación de operaciones, etc. En los libros y artículos se consideran ciertas clases de problemas, lo que es testimonio de amplia ramificación y riqueza de las investigaciones combinatorias.

La tarea de este libro consiste en exponer el análisis combinatorio como una teoría matemática de investigaciones de los conjuntos discretos en sus diferentes interpretaciones, teniendo presente que dicha teoría está basada en principios únicos, es suficientemente rica en lo que se refiere a los métodos que se emplean y puede servir de base teórica general para los métodos discretos de investigación en las matemáticas.

1.2. CONOCIMIENTOS INDISPENSABLES DE LA TEORÍA DE LOS CONJUNTOS Y DEL ÁLGEBRA

La información que se incluye en este párrafo del libro se destina para describir las propiedades de los objetos que se estudian en el análisis combinatorio, las operaciones que se realizan sobre ellos y para introducir un simbolismo uniforme. Las definiciones, los términos y los símbolos se eligen, por regla, de los que son de uso general en las matemáticas. La introducción de ellos será paulatina.

El concepto principal es el de *conjunto*. No damos definición de este concepto debido a una tautología inevitable. Todos los conjuntos que se analizan en el libro (si no hay referencias a una precisión especial) son discretos y finitos. Los conjuntos se designarán mediante letras mayúsculas latinas $A, B, \dots$, y los *elementos de los conjuntos*, mediante letras minúsculas $a, b, \dots$. Escribiremos $a \in A$, si el elemento a pertenece al conjunto A , y $a \notin A$, si a no pertenece al conjunto A . Se denomina *subconjunto* A del conjunto S ($A \subseteq S$) cualquier conjunto cuyos elementos pertenecen todos a S . En otras palabras, $A \subseteq S$, si de $a \in A$ se deduce que $a \in S$. Los conjuntos A y B *coinciden* o son *iguales* ($A = B$), si están compuestos por unos mismos elementos. Dicho de otro modo, $A = B$, cuando y sólo cuando $A \subseteq B$ y $B \subseteq A$. Si $A \subseteq B$, pero $A \neq B$, suele decirse que A es un subconjunto *propio* del conjunto B , y se escribe: $A \subset B$.

Los conjuntos, al igual que los subconjuntos, pueden introducirse y designarse mediante diferentes métodos. Por ejemplo, un subconjunto A del conjunto S se define frecuentemente como conjunto de todos los elementos $a \in S$ que poseen cierta propiedad bien determinada. Si designamos esta propiedad mediante $P(a)$, la definición de A puede escribirse simbólicamente así: $A = \{a \in S \mid P(a)\}$, o bien (simplemente) $A = \{a \mid P(a)\}$. Estas inscripciones se leen del modo siguiente: " A es un conjunto de todos los elementos

a del conjunto S , para los cuales resulta válida la condición $P(a)$ ". Por ejemplo, la inscripción $A = \{a \mid a \in \mathbb{N} \text{ y } a = 2b \text{ para cierto } b \in \mathbb{N}\}$, donde $\mathbb{N}$ es un conjunto de números naturales, describe el conjunto de todos los números positivos pares.

Introduzcamos el símbolo $\emptyset$ para designar con él un conjunto *vacio*, es decir, conjunto sin elementos. Es evidente que cualquier conjunto contiene en calidad de subconjunto un conjunto vacío. Por $P(S)$ designaremos un conjunto de todos los subconjuntos del conjunto S , el cual contiene, en calidad de elementos, todos los subconjuntos propios A que satisfacen la condición $\emptyset \subset A \subset S$, como también el conjunto vacío $\emptyset$ y el mismo conjunto S . Por ejemplo, si $S = \{a, b, c\}$, entonces $P(S)$ contiene, además de $\emptyset$ y S , seis subconjuntos propios $\{a\}$, $\{b\}$, $\{c\}$, $\{a, b\}$, $\{a, c\}$, $\{b, c\}$. En general, si S se compone de n diferentes elementos, entonces $P(S)$ contiene 2^n diferentes subconjuntos. Por eso, en la literatura matemática $P(S)$ se denota también con 2^S .

Diremos que está dada la *aplicación* φ del conjunto A en el conjunto B , si a todo elemento del conjunto A se le ha puesto en correspondencia cierto elemento del conjunto B . Para designar la aplicación φ de A en B , haremos uso del símbolo $\varphi: A \rightarrow B$. Si $a \in A$, entonces el elemento de B , que se le ha puesto en correspondencia al elemento a , se designa mediante $\varphi(a)$ y se denomina *imagen* del elemento a en la aplicación φ . Si $b \in B$, entonces cualquier elemento a de A , para el cual se verifica la igualdad $b = \varphi(a)$, se llama *preimagen* (imagen recíproca) del elemento b . El conjunto A lleva el nombre de *origen* de la aplicación φ , y el conjunto B es el *fin* de la aplicación citada. Por supuesto, cada elemento del origen de la aplicación φ tiene exactamente una sola imagen. No obstante, no todo elemento del fin de dicha aplicación tiene necesariamente una preimagen. Por otra parte, el fin de una aplicación puede contener elementos que cuentan con varias preimágenes. Un subconjunto del fin de la aplicación φ compuesto por todos los elementos suyos que tienen preimagen, o bien, lo que es lo mismo, que se representan en la forma $\varphi(a)$ para cierto $a \in A$, se denomina *imagen de la aplicación* φ y se designa $\text{Im } \varphi$. Para un conjunto finito $A = \{a_1, a_2, \dots, a_n\}$ se emplea con frecuencia la inscripción de la aplicación $\varphi: A \rightarrow B$ en forma de dos filas, a saber

$$\varphi = \begin{pmatrix} a_1 & a_2 & \dots & a_n \\ \varphi(a_1) & \varphi(a_2) & \dots & \varphi(a_n) \end{pmatrix}.$$

Dos aplicaciones $\varphi_1: A_1 \rightarrow B_1$; $\varphi_2: A_2 \rightarrow B_2$ se consideran *iguales*, si $A_1 = A_2$, $B_1 = B_2$, y $\varphi_1(a) = \varphi_2(a)$ para cualesquiera $a \in A_1$. La aplicación $\varphi': A' \rightarrow B$ es una *contracción sobre* A' de la aplicación $\varphi: A \rightarrow B$, siempre que $A' \subseteq A$ y $\varphi'(a) = \varphi(a)$ para todo $a \in A'$.

La aplicación $\varphi: A \rightarrow B$ lleva el nombre de *encaje* (inmersión), si cada elemento de B tiene no más de una preimagen, es decir, $\varphi(a) = \varphi(b)$ lleva

consigo $a = b$. Si cada elemento de B tiene al menos una preimagen, o bien, con otras palabras, $\text{Im } \varphi = B$, la aplicación φ se denomina *superposición*. Una superposición $\varphi : A \rightarrow B$ se llama también aplicación de A sobre B . Una aplicación que representa a la vez un encaje y una superposición se denomina *biunívoca* (biyectiva).

Si $A = B$, la aplicación biunívoca $\varphi : A \rightarrow A$ se llama *sustitución*.

Los conjuntos pueden ser *finitos* (es decir, compuestos por un número finito de elementos) e *infinitos*. El número de elementos en un conjunto A recibe el nombre de *potencia* del conjunto A y se designa con $|A|$. Se dice que los conjuntos A y B son *equipotentes*, si existe una aplicación biunívoca $\varphi : A \rightarrow B$. Evidentemente, dos conjuntos finitos son equipotentes, cuando y sólo cuando contienen ambos un número igual de elementos. Denominaremos *n-conjunto* a todo conjunto finito A tal que $|A| = n$. Los conjuntos infinitos pueden ser equipotentes a alguno de sus subconjuntos; por ejemplo, el conjunto de los números naturales es equipotente al conjunto de los números positivos pares y al conjunto de sus cuadrados. Entre los conjuntos infinitos se distinguen los conjuntos numerables e innumerables. Se llama *numerable* todo conjunto infinito que sea equipotente con el conjunto de números naturales. En adelante, cuando se trate de conjuntos infinitos, éstos se considerarán numerables.

Definamos ciertas operaciones que se realizan en el conjunto $P(S)$.

Se llama *unión* (reunión o suma) $A \cup B$ de los conjuntos A y B el conjunto de todos los elementos pertenecientes o bien a A , o bien a B (o bien a A y B simultáneamente):

$$A \cup B = \{a \in S \mid a \in A, \text{ o bien } a \in B\}.$$

Se llama *intersección* (o *producto*) $A \cap B$ de los conjuntos A y B el conjunto de todos los elementos pertenecientes tanto a A , como a B :

$$A \cap B = \{a \in S \mid a \in A \text{ y } a \in B\}.$$

De un modo análogo se definen las uniones e intersecciones de un sistema arbitrario (incluido el infinito) de conjuntos.

Se llama *complemento* del conjunto A el conjunto

$$\overline{A} = \{a \in S \mid a \notin A\}.$$

Se llama *diferencia* de los conjuntos A y B el conjunto $A \setminus B = \{a \in S \mid a \in A \text{ y } a \notin B\}$. Es evidente que $A \setminus B = A \cap \overline{B}$.

Se llama *diferencia simétrica* de los conjuntos A y B el conjunto

$$A \Delta B = (A \cup B) \setminus (A \cap B) = (A \setminus B) \cup (B \setminus A).$$

El conjunto $P(S)$ con las operaciones de unión, intersección y de complemento definidas sobre él recibe el nombre de *booleana*. Quedan válidas en este caso las siguientes leyes algebraicas:

1. $A \cap A = A$; $A \cup A = A$ (idempotencia);
2. $A \cap B = B \cap A$; $A \cup B = B \cup A$ (conmutatividad);
3. $A \cap (B \cap C) = (A \cap B) \cap C$; $A \cup (B \cup C) = (A \cup B) \cup C$ (asociatividad);
4. $A \cap (A \cup B) = A$; $A \cup (A \cap B) = A$ (absorción);
5. Si $A \subseteq C$, se tiene $A \cup (B \cap C) = (A \cup B) \cap C$ (modularidad);
6. $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$;
 $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$ (distributividad);
7. $A \cap \emptyset = \emptyset$; $A \cup \emptyset = A$;
 $A \cap \overline{S} = \overline{A}$; $A \cup \overline{S} = \overline{S}$ (fronteras universales);
8. $A \cap \overline{A} = \emptyset$; $A \cup \overline{A} = S$ (complementariedad);
9. $\overline{\overline{A}} = A$
10. $\overline{A \cap B} = \overline{A} \cup \overline{B}$; $\overline{A \cup B} = \overline{A} \cap \overline{B}$ (leyes de De Morgan).

Si $A \cap B = \emptyset$, entonces A y B se llaman conjuntos *disjuntos*. La representación del conjunto S en forma de una unión de los subconjuntos disjuntos dos a dos recibe el nombre de *partición* del conjunto S . En otras palabras, $A_1, A_2, \dots, A_k$ es una partición del conjunto S , si $S = \bigcup_{i=1}^k A_i$ y

$A_i \cap A_j = \emptyset$ para cualesquiera $i \neq j$. Una totalidad de todos los pares ordenados (a, b) tales que $a \in A$, $b \in B$, se denomina producto *directo* (*cartesiano*) de los conjuntos A y B y se designa por $A \times B$, es decir,

$$A \times B = \{(a, b) \mid a \in A, b \in B\}.$$

Si $A = B$, el producto directo $A \times A$ se denota con A^2 . Análogamente se determina el producto directo de k conjuntos:

$$A_1 \times A_2 \times \dots \times A_k = \{(a_1, a_2, \dots, a_k) \mid a_1 \in A_1, a_2 \in A_2, \dots, a_k \in A_k\},$$

y, al coincidir los conjuntos, tenemos: $A^k = A \times A \times \dots \times A$.

Enunciemos las reglas evidentes que yacen en la base de muchos cálculos y estimaciones combinatorios.

Regla de la suma. Si S es un conjunto finito y $S = \bigcup_{i=1}^k A_i$, donde $A_i \in P(S)$, $i = 1, 2, \dots, k$, entonces

$$|S| = \left| \bigcup_{i=1}^k A_i \right| \leq \sum_{i=1}^k |A_i|,$$

con la particularidad de que la igualdad tiene lugar cuando $A_1, A_2, \dots, A_k$ forman una partición del conjunto S .

Regla del producto. Para los conjuntos finitos $A_1, A_2, \dots, A_k$ tenemos

$$|A_1 \times A_2 \times \dots \times A_k| = \prod_{i=1}^k |A_i|.$$

Llamemos *correspondencia binaria* entre los conjuntos A y B cualquier subconjunto $A \subseteq A \times B$. Si $(a, b) \in \alpha$, suele decirse que a se encuentra en

Fig.1.1



relación α respecto a b , y esto se expresa en la forma $a\alpha b$. La noción de correspondencia binaria entre A y B sirve de generalización del concepto de la aplicación $\varphi : A \rightarrow B$. En efecto, cada aplicación $\varphi : A \rightarrow B$ define la correspondencia binaria α_φ entre A y B : la expresión $a\alpha_\varphi b$ significa que $b = \varphi(a)$. Viceversa, sea dada una correspondencia binaria α entre los conjuntos A y B . Analicemos para todo $a \in A$ el conjunto de todos los $b \in B$ que se caracteriza por la propiedad $a\alpha b$. Dicha correspondencia define la aplicación $\varphi : A \rightarrow B$, cuando y sólo cuando para todo $a \in A$ existe exactamente un solo elemento $b \in B$ que se caracteriza por las propiedades $a\alpha b$. Así pues, la noción de correspondencia binaria incluye el concepto de aplicación como un caso particular (muy importante).

Cuando $A = B$, la correspondencia binaria $q \subseteq A \times A$ lleva el nombre de *relación binaria* en el conjunto A . Son ejemplos de relación binaria la relación de igualdad $\Delta = \{(a, a) \mid a \in A\}$, llamada *diagonal* del conjunto A , y la relación que coincide con todo el conjunto $A \times A$, denominada *relación unidad*.

A una relación binaria q sobre el conjunto finito A se le puede poner en correspondencia un ente geométrico llamado grafo orientado o diagrama. A todo elemento $a \in A$ se le pone en correspondencia un punto en el plano que lleva el nombre de vértice. Si para a y b tenemos aqb , los puntos marcados con a y b se unen por una flecha que parte de a y llega a b y que se denomina arco. La totalidad de vértices y arcos construidos del modo mencionado representan el diagrama $\Gamma(A, q)$ de la relación q . En la fig. 1.1 se exponen los diagramas $\Gamma(A, \Delta)$ y $\Gamma(A, I)$, respectivamente, donde $A = \{a, b, c\}$, $I = A \times A$.

Sea A' un subconjunto no vacío del conjunto A , sobre el cual viene dada la relación binaria q : de la relación q' en A' , definida mediante la condición

$$aq'b = aqb \text{ para cualesquiera } a, b \in A',$$

se dice que ella está *inducida* mediante la relación q .

Formulemos las propiedades que pueden tener las relaciones binarias. Una relación q se denomina *reflexiva*, si aga para todo $a \in A$. Dicho de otro modo, una relación es reflexiva, siempre que $\Delta \subseteq q$. Si la expresión aqb lleva consigo bqa , la relación q se llama *simétrica*. En cambio, si aqb y bqa tienen por resultado $a = b$, la relación q se llama *antisimétrica*. Una relación q se llama *transitiva*, si aqb y bqc conducen a aqc . No es difícil

convencerse de que la diagonal posee todas las cuatro propiedades citadas, mientras que la relación unidad tiene todas estas propiedades, a excepción de la antisimetría. Pero, si A se compone sólo de un elemento, la relación unidad coincide con la diagonal.

Se denomina *clase contigua* de la relación binaria q , definida por un elemento $a \in A$, el conjunto de todos los elementos $b \in A$ tales que bqa .

La relación binaria q en el conjunto A se llama relación de equivalencia, si es reflexiva, simétrica y transitiva a la vez. Las relaciones de equivalencia son de gran importancia en el análisis combinatorio, puesto que están estrechamente ligadas con las particiones de los conjuntos. Efectivamente, puede demostrarse con facilidad que una totalidad de clases contiguas de relaciones de equivalencia sobre un conjunto S representa la partición de este último y que a toda partición del conjunto S le corresponde la relación de equivalencia q , cuyas clases coinciden con los bloques de la partición indicada (la afirmación aqb tiene lugar, si y sólo si a y b pertenecen a un mismo subconjunto de la partición del conjunto S).

Una relación binaria q sobre el conjunto A se denomina relación de orden parcial, si es reflexiva, antisimétrica y transitiva. La relación de orden parcial se denotará con el símbolo $\leq$. La notación $a \geq b$ significa que $b \leq a$. A veces diremos que A (o bien $(A, \leq)$) es un conjunto parcialmente ordenado, suponiendo que la relación de orden parcial ya es conocida (esta afirmación, sin embargo, no es bien correcta, por cuanto en un solo conjunto pueden fijarse órdenes diferentes).

Un conjunto parcialmente ordenado se denomina *trivial*, siempre que $a \leq b$, cuando y sólo cuando $a = b$. Un conjunto parcialmente ordenado $(A, \leq)$, que posee la propiedad de que para cualesquiera $a, b \in A$ se verifica o bien $a \leq b$, o bien $b \leq a$, lleva el nombre de *cadena* (lo llaman también conjunto *linealmente ordenado*).

Recordemos, además, algunos conceptos algebraicos que nos harán falta ulteriormente. Se denomina *operación binaria* sobre el conjunto A la aplicación $f: A \times A \rightarrow A$. Si $f(a, b) = c$, donde $a, b, c \in A$, se escribe $a * b = c$. Una operación binaria se llama *asociativa*, si para cualesquiera $a, b, c \in A$:

$$(a * b) * c = a * (b * c).$$

Si una operación binaria satisface la condición

$$a * b = b * a$$

para cualesquiera $a, b \in A$, se denomina *conmutativa*. El elemento $e \in A$ se denomina elemento *unidad* (elemento *neutro*) respecto de la operación $*$, si para todo $a \in A$:

$$a * e = e * a = a.$$

Si tal elemento existe, es único. Supongamos que un conjunto A con operación prefijada posee un elemento unidad e . En este caso el elemento a^{-1}

se denomina *inverso* (o *simétrico*) del elemento $a \in A$ con relación a dicha operación, siempre que $a * a^{-1} = a^{-1} * a = e$. El conjunto A con operación binaria $*$, en el cual las ecuaciones $a * x = b$, e $y * a = b$ son resolubles unívocamente respecto de x e y para cualesquiera $a, b \in A$ se llama *casi grupo*. Un conjunto no vacío A , sobre el cual viene definida la operación binaria asociativa, recibe el nombre de *semigrupo*. Si la operación es conmutativa, el semigrupo se llama *abeliano*. Un semigrupo, en el que existe un elemento unidad y para cada elemento existe uno inverso, se denomina *grupo*. El número de elementos en el grupo se llama *orden* del grupo. Un grupo cuyos elementos son todos la potencia de un elemento a (es decir, pueden obtenerse por la aplicación sucesiva de la operación, empleando un mismo elemento a distinto del elemento neutro) se llama *cíclico*. Los grupos cíclicos son siempre abelianos.

Sean A y B los grupos finitos de órdenes arbitrarios con las operaciones binarias $*$ y $\circ$, respectivamente, y supongamos que $\varphi: A \rightarrow B$ es una aplicación, con la cual para cualesquiera $a, a' \in A$ se verifica la igualdad

$$\varphi(a * a') = \varphi(a) \circ \varphi(a'),$$

Tal aplicación φ se llama *homomorfismo* del grupo A en el grupo B . Si φ es una aplicación biunívoca, el homomorfismo recibe el nombre de *isomorfismo*.

Ejemplo. Un conjunto de sustituciones $\varphi: X \rightarrow X$ forma un grupo, si la operación $*$ se define como resultado de la actuación sucesiva de dos sustituciones, es decir $\varphi * \psi = \pi$, donde $\forall x \in X: \pi(x) = \psi(\varphi(x))$. Dicho grupo lleva el nombre de *grupo simétrico de potencia $n = |X|$* . Es fácil mostrar que todos los grupos simétricos de un mismo orden son isomorfos. Observemos que los grupos simétricos no son abelianos cuando $n \geq 3$.

Si en un subconjunto $B \subseteq A$ el resultado de una operación de dos cualesquiera elementos de B pertenece también a B , suele decirse que B está *cerrado* respecto de la operación dada. Se llama *subgrupo* a un subconjunto no vacío del grupo que está cerrado respecto de la operación binaria y que para cada elemento contiene un elemento inverso.

Un conjunto A con dos operaciones binarias $+$ y $\circ$ se denomina *anillo*, siempre que forma un grupo abeliano respecto de la operación $+$, un semigrupo respecto de la operación $\circ$, y si $\circ$ es distributiva respecto de la operación $+$, es decir, si para cualesquiera $a, b, c \in A$ se tiene

$$\begin{aligned}(a + b) \circ c &= a \circ c + b \circ c, \\ c \circ (a + b) &= c \circ a + c \circ b.\end{aligned}$$

En un anillo los elementos neutros respecto de las operaciones $+$ y $\circ$ se denominan *nulo* y *unidad*, respectivamente. Un elemento del anillo se llama *invertible*, si para dicho elemento existe un elemento inverso respecto de la operación $\circ$.

Si un conjunto de elementos de un anillo, distintos de cero, forma res-

pecto de la operación $\circ$ un grupo abeliano, entonces el anillo se llama en este caso *campo*. Un campo en el que se contienen un número finito de elementos se llama *finito*. Un campo finito de n elementos existe, cuando y sólo cuando $n = p^\alpha$, donde p es un número primo y α , un número natural. Tal campo es único con exactitud hasta un isomorfismo en el que se conservan ambas operaciones binarias. Este campo se denomina *de Galois* y se designa, de ordinario, con la expresión $GF(p^\alpha)$. Si $\alpha = 1$, $GF(p)$ es isomorfo al campo de residuos respecto al módulo primo (o irreducible) p .

Como ejemplos de un anillo infinito y de un campo infinito sirven el anillo de números enteros $\mathbb{Z}$ y el campo de números reales $\mathbb{R}$. Las operaciones binarias son aquí operaciones corrientes de adición y multiplicación.

Se denomina *espacio lineal* L sobre el campo P el conjunto dotado de la operación binaria $\varphi: L \times L \rightarrow L$, la cual se designa habitualmente como adición $\varphi(a, b) = a + b$ para $a, b \in L$, y de la operación binaria externa $f: P \times L \rightarrow L$, la cual se designa habitualmente como multiplicación $f(p, a) = pa$, con la particularidad de que dichas operaciones satisfacen los siguientes axiomas:

a) con respecto a la operación de adición el espacio L es un grupo abeliano. El elemento neutro de este grupo se denota con 0; el elemento inverso de a se designa corrientemente con $-a$;

b) la multiplicación de los vectores por los elementos del campo P , o por los escalares, es unitaria, es decir, $1 \cdot a = a$ para todos los a , y asociativa, es decir, $p(qa) = (pq)a$ para cualesquiera $p, q \in P$; $a \in L$;

c) la adición y la multiplicación están unidas mediante las leyes de distributividad, es decir,

$$p(a + b) = pa + pb,$$

$$(p + q)a = pa + qa$$

para cualesquiera $p, q \in P$; $a, b \in L$.

La expresión del tipo $p_1a_1 + \dots + p_na_n$ lleva el nombre de *combinación lineal* de los vectores $a_1, a_2, \dots, a_n$; los escalares p_i se llaman *coeficientes* de dicha combinación lineal.

Se denomina *álgebra* sobre el campo P el anillo asociativo con la unidad A que contiene el campo P y es tal que P yace en el centro de A , es decir, que conmuta con todos los elementos de A . En particular, A es un espacio lineal sobre el campo P .

Los conceptos nuevos se introducirán más adelante a medida que surja la necesidad. Paseinos, ahora, a la explicación del sentido de las operaciones combinatorias más simples. Nos limitaremos, para empezar, al análisis de los conjuntos linealmente ordenados. Sin embargo, estudiaremos también, en forma paulatina, los conjuntos de estructura más compleja.

1.3. MUESTRAS Y ORDENACIONES

Con el concepto de *muestra* se asocian tanto la propia operación de selección de un subconjunto del conjunto dado, como también el resultado de la operación citada, es decir, el subconjunto elegido. En lo sucesivo se tendrá en cuenta precisamente la segunda interpretación, siempre que no se diga lo contrario.

Supongamos que de un n -conjunto A_n se ha obtenido una r -muestra:

$(a_1, a_2, \dots, a_r)$, donde $a_i \in A_n$; $i = 1, 2, \dots, r$; $r \leq n$.

El número r se llama *volumen* de la muestra.

Según sean las condiciones del problema, en las muestras puede tomarse en consideración el orden de sucesión de los elementos en ellas (y en este caso las r -muestras se llaman *r -permutaciones*), o bien dicho orden no se toma en consideración (en este último caso se denominan *r -combinaciones*). Por ejemplo, dos 5-muestras del conjunto A_n ($n \geq 5$):

$(a_1, a_2, a_3, a_4, a_5)$ y $(a_5, a_4, a_3, a_2, a_1)$

representan en sí 5-combinaciones iguales y al mismo tiempo 5-permutaciones diferentes. En general, dos r -permutaciones $a = (a_1, a_2, \dots, a_r)$ y $b = (b_1, b_2, \dots, b_r)$ son iguales: $a = b$ únicamente si $a_i = b_i$; $i = 1, 2, \dots, r$. En las muestras es posible la aparición reiterada de los elementos, y en tal caso ellas se denominan *r -combinaciones con repetición* y *r -permutaciones con repetición*, respectivamente. Una r -permutación (con repetición) de elementos del conjunto A se llama también *palabra de longitud r sobre el alfabeto A* .

Es evidente que los conceptos de r -permutaciones y r -combinaciones, al igual que sus combinaciones, abarcan todos los tipos posibles de muestras. Por eso no hay necesidad de dar el concepto de arreglo (variación), aunque dicho concepto aparece todavía en la literatura, principalmente en los manuales.

La multiformidad de la solución de los problemas combinatorios observada en las etapas iniciales del desarrollo de las matemáticas condujo a una cuestión natural: ¿cuántos son los procedimientos, por medio de los cuales puede realizarse la requerida disposición combinatoria? En particular, el cálculo del número de r -muestras de un n -conjunto fue históricamente uno de los primeros problemas de la combinatoria.

Hallemos el número de todas las r -permutaciones posibles (sin repetición) de un n -conjunto. Denotemos el número que se busca mediante $P(n, r)$. El problema se reduce a una aplicación sucesiva de la regla del producto. En efecto, en el n -conjunto se tienen n posibilidades para elegir el primer elemento de la r -permutación. Una vez realizada tal elección, quedan $n - 1$ posibilidades para la elección del segundo elemento, luego quedan $n - 2$ posibilidades para elegir el tercer elemento, etc.: para la elección del r -ésimo

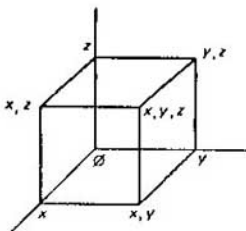


Fig 1.2.

elemento tendremos $n - r + 1$ posibilidades. De acuerdo con la regla del producto,

$$P(n, r) = n(n - 1) \dots (n - r + 1),$$

de donde se deduce

$$P(n, n) = n!$$

Para que el resultado sea más completo, admitamos

$$P(n, 0) = 0! = 1.$$

Calculemos ahora el número de r -permutaciones posibles con repetición. En este caso, después de elegir cualquier elemento de la r -permutación quedan las mismas n posibilidades para elegir el elemento siguiente. Por consiguiente, según la regla del producto, el número de r -permutaciones con repetición del n -conjunto es igual a n^r .

Los razonamientos aducidos aquí se ilustran fácilmente con un ejemplo del esquema de urna, cuyos diferentes tipos se emplean en la teoría de las probabilidades: se tiene una urna dentro de la cual se encuentran colocadas n bolas iguales y de la cual se sacan por turno r bolas. En tal caso resultan posibles dos casos: la bola sacada o bien se retorna a la urna (elección con retorno) o bien no se retorna (elección sin retorno).

Un ejemplo más. ¿Cuántos subconjuntos tiene un n -conjunto S , es decir, a qué es igual $|P(S)|$? La respuesta será $|P(S)| = 2^n$. Efectivamente, cualquier r -muestra $R = (s_{i1}, s_{i2}, \dots, s_{ir})$, donde $r = 1, 2, \dots, n$, figura en $P(S)$. A esta r -muestra se le puede poner en correspondencia una n -muestra, compuesta por elementos de dos tipos: ceros, si el elemento no integra la r -muestra R , y unidades, si el elemento figura en R . De este modo, las unidades deben disponerse en los lugares correspondientes a $i_1, i_2, \dots, i_r$, mientras que los ceros, en los demás lugares. Pero, el número de tales n -muestras (es decir, de n -permutaciones con repetición) de 2-conjunto $\{0, 1\}$ es igual a 2^n , lo que constituye precisamente el resultado buscado. Este mismo problema puede interpretarse como el problema sobre el número de vértices de un hipercubo en el espacio de n dimensiones (el caso de $n = 3$, $S = \{x, y, z\}$ se muestra en la fig. 1.2, donde todos los subconjuntos $P(S)$ son vértices del cubo).

Ejercicio. ¿Cuál es el número de matrices de dimensión $k \times l$, compuestas por ceros y unidades?

Calculemos ahora el número de r -combinaciones, designándolo con $\binom{n}{r}$ o con C_n^r . Comencemos por el caso en que todos los elementos en las r -combinaciones son diferentes. Es fácil ver que el número de r -combinaciones del n -conjunto es $r!$ veces inferior al número de r -permutaciones de los elementos del mismo conjunto. Por consiguiente,

$$\binom{n}{r} = \frac{P(n, r)}{r!} = \frac{n(n-1) \dots (n-r+1)}{r!} = \frac{n!}{r!(n-r)!};$$

de aquí se deduce que $\binom{n}{r} = \binom{n-r}{r}$; en particular,

$$\binom{n}{n} = \binom{n}{0} = 1.$$

Observemos que las r -combinaciones del n -conjunto son sus r -subconjuntos. Con este motivo estudiemos el problema sobre el número de $(r_1, r_2, \dots, r_k)$ -particiones del n -conjunto S , es decir, de particiones ordenadas del tipo $S = T_1 \cup T_2 \cup \dots \cup T_k$, en las cuales $T_i \cap T_j = \emptyset$ para $i \neq j$; $i, j = 1, 2, \dots, k$, con la particularidad de que T_i es un r_i -subconjunto del conjunto S , $i = 1, 2, \dots, k$. Obviamente, $\sum_{i=1}^k r_i = n$. Razonemos igual que lo hemos hecho al buscar el número $P(n, r)$. Para elegir el r_1 -subconjunto T_1 de S se tienen $\binom{n}{r_1}$ posibilidades: entonces el r_2 -subconjunto T_2 puede elegirse sólo de $n - r_1$ elementos restantes (ya que $T_1 \cap T_2 = \emptyset$), y, por lo tanto, se tienen $\binom{n-r_1}{r_2}$ métodos para la elección de T_2 , etc.; el r_k -subconjunto T_k puede ser elegido sólo después de haber elegido los r_i -conjuntos T_i , $i = 1, 2, \dots, k-1$, por consiguiente de $n - \sum_{i=1}^{k-1} r_i$ elementos restantes

puede elegirse, sirviéndose de $\binom{n - \sum_{i=1}^{k-1} r_i}{r_k}$ métodos. Aplicando ahora

la regla del producto, obtendremos que el número buscado de $(r_1, r_2, \dots, r_k)$ -particiones del n -conjunto S es igual a

$$R = \binom{n}{r_1} \binom{n-r_1}{r_2} \dots \binom{n - \sum_{i=1}^{k-1} r_i}{r_k} = \frac{n!}{r_1! r_2! \dots r_k!},$$

(tomando en consideración la expresión para $\binom{n}{r}$). La r -combinación

del n -conjunto puede ser interpretada como una $(r, n - r)$ -partición, y la $(1, 1, \dots, 1)$ -partición representa simplemente una permutación. Calculemos, por fin, el número de r -combinaciones con repetición del n -conjunto S . Daremos a conocer tres métodos distintos de obtención de este número con el objeto de mostrar los rasgos específicos de los razonamientos combinatorios.

1^{er} método. Admitamos que los elementos del conjunto S están numerados por medio de los números $1, 2, \dots, n$ (es decir, S se encuentra en una correspondencia biunívoca con el conjunto de los primeros n números naturales). Entonces, en lugar de las r -muestras del conjunto S podemos analizar las r -muestras (biunívocas) del conjunto $S' = \{1, 2, \dots, n\}$ que corresponden a las primeras. Toda r -muestra del conjunto S' puede ser escrita en la forma $A = (a_1, a_2, \dots, a_r)$, donde $a_1 \leq a_2 \leq \dots \leq a_r$ (la igualdad de los números responde al caso de elementos iguales en la r -muestra correspondiente de S). A la r -muestra A (los elementos en esta muestra no son forzosamente diferentes) le ponemos en correspondencia el r -conjunto $A' = \{a_1 + 0, a_2 + 1, \dots, a_r + r - 1\}$, en el cual todos los elementos son, evidentemente, diferentes.

Como es fácil de ver, dicha correspondencia es biunívoca, con la particularidad de que los r -conjuntos A' son r -combinaciones sin repetición del $(n + r - 1)$ -conjunto $\{1, 2, \dots, n, n + 1, \dots, n + r - 1\}$, cuyo número es igual (como quedó demostrado) a $\binom{n + r - 1}{r}$. Esta es precisamente la respuesta que buscábamos.

El 2^o método consiste en la obtención de una fórmula recurrente¹⁾.

Denotemos mediante $f(n, r)$ el número de r -combinaciones con repeticiones del n -conjunto S . Está claro que

$$f(n, 1) = n \quad \text{y} \quad f(1, r) = 1.$$

(para cualquier $n > 0$ entero de n elementos pueden elegirse n diferentes 1-combinaciones, es decir, n diferentes elementos; mientras tanto para cualquier $r > 0$ entero de un elemento se puede obtener solamente una r -combinación: una r -muestra compuesta por r elementos iguales). Fijemos en S cierto elemento, entonces cada r -combinación o bien contiene este elemento o no lo contiene. Si tiene lugar el primer caso, los demás $r - 1$ elementos de esta r -combinación (y, por tanto, las r -combinaciones que contienen el elemento fijado) pueden ser elegidos empleando los $f(n, r - 1)$ métodos. Si tiene lugar el segundo caso, la r -combinación se elige de $n - 1$ elementos, y entonces, el número de tales r -combinaciones es igual a $f(n - 1, r)$. Empleando la regla de la suma, obtendremos

¹⁾ Se denominan fórmulas (relaciones) recurrentes aquellas que permiten calcular los valores de una magnitud buscada paso a paso, partiendo de los valores «iniciales» conocidos y de los calculados de antemano.

$$f(n, r) = f(n, r-1) + f(n-1, r). \quad (1)$$

En particular, al conocer $f(n, 1)$ y $f(1, r)$, tenemos

$$f(n, 0) = f(n, 1) - f(n-1, 1) = 1,$$

lo que concuerda con el resultado obtenido anteriormente. Ahora obtenemos sucesivamente

$$f(n, 2) = f(n, 1) + f(n-1, 2) = f(n, 1) + f(n-1, 1) + f(n-2, 2) = \dots$$

$$\dots = n + (n-1) + (n-2) + \dots + 1 = \frac{n(n+1)}{2} = \binom{n+1}{2};$$

$$\begin{aligned} f(n, 3) &= f(n, 2) + f(n-1, 2) + \dots + f(1, 3) = \\ &= \binom{n+1}{2} + \binom{n}{2} + \dots + 1 = \binom{n+2}{3} \end{aligned}$$

etc. Es fácil convencerse de que

$$f(n, r) = \binom{n+r-1}{r}$$

satisface la correlación (1) y las condiciones iniciales:

$$f(n, 1) = n; f(1, r) = 1.$$

3er método. A una r -combinación con repetición del n -conjunto S (por ejemplo, a bcb de $S = \{a, b, c, d, e\}$) agreguemos todos los n elementos del conjunto S , y escribamos por orden los $n+r$ elementos obtenidos, disponiendo juntos los elementos iguales: ($abbbccde$). A continuación, dividamos los subconjuntos de elementos iguales mediante $n-1$ rayas: ($a | bbb | b | cc | d | e$). Por fin, sustituyamos por puntos todos los elementos dispuestos entre las rayas: ($\cdot | \dots | \cdot | \cdot | \cdot$). De este modo, a la r -combinación se le asigna (equipara) la colocación de $n-1$ rayas en $n+r-1$ intervalos entre $n+r$ puntos. Inversamente en cada colocación de esta índole se restaura inóvamente la r -combinación que le corresponde. Por ejemplo,

$$(\cdot | \cdot | \cdot | \cdot | \cdot | \cdot) \rightarrow (aa | b | cc | d | ee) \rightarrow (aahccdeee) \rightarrow \{a, c, e\}.$$

Existen en total $\binom{n+r-1}{n-1} = \binom{n+r-1}{r}$ métodos de colocar $n-1$ rayas en $n+r-1$ lugares. Por consecuencia, existe exactamente la misma cantidad de r -combinaciones con repetición del n -conjunto.

Como conclusión de este párrafo, analicemos un concepto relacionado con la operación de ordenación, es decir, la permutación. Esta última puede ser examinada desde dos puntos de vista: a) como una totalidad ordenada de elementos del conjunto dado, o bien b) como una perturbación del orden estándar llamado, habitualmente, natural (por ejemplo, alfabético, numérico). El caso a) conduce a las r -permutaciones, $r \leq n$, ya descritas anteriormente. El caso b) conduce a las n -permutaciones (en relación con la defini-

ción de la r -permutación) denominadas simplemente permutaciones (o sustituciones), que se estudian en la teoría de los grupos.

Sea, por ejemplo una permutación

$$P = (4, 3, 7, 5, 6, 9, 2, 8, 1, 12, 11, 10),$$

que representa una perturbación del orden natural de los primeros 12 números de la serie natural. Puede ser escrita en forma de una sustitución (en la primera fila se pone el orden natural y en la segunda, el perturbado).

$$P = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 \\ 4 & 3 & 7 & 5 & 6 & 9 & 2 & 8 & 1 & 12 & 11 & 10 \end{pmatrix}.$$

Esta notación muestra que, al realizarse la permutación P , el elemento 1 se convierte en 4, el elemento 2 en 3, el 3 en 7, etc. La permutación P puede escribirse también de otro modo:

$$P = (1, 4, 5, 6, 9)(2, 3, 7)(10, 12)(8)(11), \quad (2)$$

donde cada paréntesis es una permutación que actúa sólo contra los elementos encerrados dentro del paréntesis dado y no toca los elementos no encerrados en él (por ejemplo, la sustitución $(2, 3, 7)$ convierte 2 en 3, 3 en 7, y 7, de nuevo en 2). La representación de una permutación en la forma (2) lleva el nombre de descomposición *en ciclos*. Cualquier permutación puede ser descompuesta en ciclos. Esta descomposición es única, con una exactitud de hasta las permutaciones cíclicas de los elementos dentro de los ciclos. Por ejemplo, $(2, 3, 7)$, $(3, 7, 2)$ $(7, 2, 3)$ son notaciones diferentes de un mismo ciclo.

Supongamos que una permutación contiene k_1 ciclos compuestos por un solo elemento, es decir, 1-ciclos, luego k_2 2-ciclos, k_3 3-ciclos, etc. En este caso se denomina $(k_1, k_2, \dots, k_n)$ -permutación, o bien permutación de la forma

$$1^{k_1} 2^{k_2} \dots n^{k_n}, \quad (3)$$

donde, evidentemente,

$$\sum_{i=1}^n i k_i = n.$$

Teorema. El número de permutaciones del tipo (3) es igual a

$$P(k_1, k_2, \dots, k_n) = \frac{n!}{1^{k_1} k_1! 2^{k_2} k_2! \dots n^{k_n} k_n!}. \quad (4)$$

Demostración. Examinemos la notación de una descomposición en ciclos para la permutación del tipo (3), a saber: primero k_1 paréntesis para la notación de los ciclos de longitud 1, luego k_2 paréntesis para la notación de los ciclos de longitud 2, etc. En n posiciones dispuestas dentro de todos los paréntesis podemos poner n elementos, sirviéndonos de $n!$ métodos, y cada vez obtendremos la notación de la permutación del tipo (3). Sin em-

bargo, entre dichas $n!$ notaciones se encuentran diferentes notaciones de una misma permutación. Aclaremos cuántas notaciones diferentes tiene una permutación. En primer lugar como hemos observado anteriormente, cada ciclo de longitud i puede escribirse dentro de los márgenes del paréntesis dado mediante i métodos. En segundo lugar, sirviéndose de k_i métodos se pueden reordenar los paréntesis donde están escritos los ciclos de longitud i . Según la regla del producto obtenemos, que una familia de ciclos de longitud i puede ser representada por $i^{k_i} \cdot k_i!$ métodos. Al hacer i recorrer los valores de 1 a n , y al aplicar otra vez la regla del producto, concluimos que existen

$$N = 1^{k_1} k_1! 2^{k_2} k_2! \dots n^{k_n} k_n!$$

métodos para escribir cada permutación del tipo (3). Por consiguiente, se tienen en total $n!/N$ de tales permutaciones, lo que era necesario demostrar.

La representación de las permutaciones en forma de un producto de ciclos sirve de fuente para muchos problemas combinatorios, por ejemplo: hallar el número de permutaciones del n -conjunto que tengan un número prefijado de ciclos (sin tomar en consideración la longitud de los ciclos); que dejen los elementos dados inmóviles; que tengan un número dado de ciclos de longitud prefijada, etc.

1.4. DISTRIBUCIONES Y LLENADOS

En muchos problemas cierta totalidad de elementos (por ejemplo de granos, pernos, tuercas, etc.) se distribuye por cierto conjunto de células (cajas, cajones, etc.) que, como consecuencia, se llenan. Ambos conceptos principales, esto es, distribución (su sinónimo es la partición) y llenado, se emplean tanto para designar las operaciones, como para expresar el resultado de ellas, es decir, la situación obtenida.

Los problemas de esta clase existen desde hace mucho tiempo y cuentan con un procedimiento de resolución elaborado. El interés hacia estos problemas no se extingue, pues tienen gran importancia práctica. Ellos se manifiestan en los más diversos planteamientos: particiones de conjuntos, cortes de grafos y de redes, agrupaciones de tornos y de mecanismos automáticos y robots, etc.

En el aspecto teórico dichos problemas pueden interpretarse como aplicaciones de un conjunto (de elementos) sobre otro conjunto (de células). Pueden ser tratados también como elección de muestras.

Los medios adoptados de resolución de esta clase de problemas dependen de las condiciones impuestas sobre los tipos de elementos a distribuir, métodos de distribución y capacidad de las células. Es evidente, que la riqueza de las condiciones posibles determina la diversidad de los procedimientos aplicados para la resolución de los problemas. Más abajo se da cierta información que introduce al lector en esta esfera de problemas.

Para el cálculo del número de distribuciones hace falta precisar si los ele-

mentos del conjunto dado y las células son distinguibles (por ejemplo, numerados) o no. De acuerdo con ello los problemas se dividen en las siguientes cuatro clases.

(A) los elementos del conjunto son distinguibles, como también lo son las células;

(B) los elementos del conjunto no son distinguibles, las células, son distinguibles;

(C) los elementos del conjunto son distinguibles, las células no son distinguibles;

(D) tanto los elementos del conjunto, como las células no se distinguen entre sí.

Dentro de cada una de estas clases los problemas se diferencian, a su vez, por la forma de las aplicaciones prefijadas por las condiciones concretas. Convengamos en que a continuación N siempre servirá de designación para un n -conjunto de elementos, y R será un r -conjunto de células. Por cuanto en este párrafo hablamos sólo de los fundamentos teóricos de la operación de distribución y llenado, estudiemos estas clases de problemas en rasgos generales, sin tratar de exponer completa y detalladamente todos los accesos a la resolución de los problemas correspondientes.

(A) Según lo dicho anteriormente, todos los elementos del conjunto N y todas las células del conjunto R son distinguibles. Aquí no tiene importancia si se diferencian por la forma, el color, volumen o, incluso, por el número. Tiene importancia sólo el hecho de distinción. He aquí algunas formas equivalentes de este problema: a) formación de las palabras de longitud r , a partir del alfabeto compuesto por n letras; b) extracción sucesiva de r bolas de una urna que contiene n bolas, con su retorno inmediato; c) formación de las r -permutaciones con repetición de n símbolos.

El carácter de la aplicación que por ahora está libre de limitaciones algunas, permite indicar en seguida el número de distribuciones posibles:

$$P = n^r,$$

por cuanto para cada una de r células se tiene la posibilidad de colocar en ella cualquiera de n elementos. La forma particular de las aplicaciones (biunívocas) corresponde a una limitación adicional: en cada célula cabe uno, y sólo un elemento. En este caso

$$P = n(n-1) \dots (n-r+1) = \frac{n!}{(n-r)!}.$$

A la clase (A) pertenecen, en particular, los siguientes casos de distinción de los elementos y de las células.

(A₁) El conjunto N tiene la $(k_1, k_2, \dots, k_m)$ -especificación, si cuenta con k_1 elementos del primer tipo (por ejemplo, de color), k_2 elementos del segundo tipo, ..., k_m elementos del m -ésimo tipo (además, $k_1 + k_2 + \dots + k_m = n$).

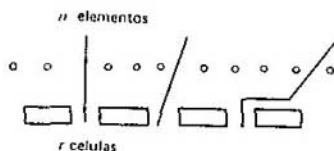


Fig.1.3.

(A₂) El r -conjunto R tiene la $(n_1, n_2, \dots, n_r)$ -especificación, si en la i -ésima célula se disponen n_i elementos, $i = 1, 2, \dots, r$.

(A₃) Los elementos dentro de las células están ordenados, es decir, dos células se consideran llenas de una manera diferente, si es distinta la ordenación de los elementos (incluso de los mismos) alojados en ellas; no hay limitaciones para el volumen de las células.

Sin pretender dar una descripción completa de todas las situaciones posibles, aduzcamos un solo ejemplo. Supongamos, por ejemplo, que el conjunto N tiene una (p, q) -especificación, es decir, contiene p elementos del primer tipo y q elementos del segundo tipo; $p + q = n$. Se necesita determinar cuántas son las distribuciones de los elementos del conjunto N en r células diferentes sin limitaciones para el número de elementos en cada una de las células. Los elementos del primer tipo pueden distribuirse en r células por

$$\binom{p + r - 1}{r - 1} = \binom{p + r - 1}{p}$$

métodos, y los elementos del segundo tipo, por

$$\binom{q + r - 1}{r - 1} = \binom{q + r - 1}{q}$$

métodos. El número total de distribuciones es igual, en virtud de la regla del producto, a

$$\binom{p + r - 1}{p} = \binom{q + r - 1}{q}.$$

Si tiene lugar la $(k_1, k_2, \dots, k_m)$ -especificación del conjunto N , el número de distribuciones de n elementos suyos en r células diferentes será

$$\prod_{i=1}^m \binom{r + k_i - 1}{k_i}.$$

Pasemos a los problemas de la clase (B). Como se ha dicho, en los problemas de este tipo los elementos del conjunto N no son distinguibles y los del conjunto R , distinguibles. Examinemos diferentes casos:

1. Los elementos del conjunto N están distribuidos en las células del conjunto R de una manera tal que ninguna célula está vacía (en la fig. 1.3. se muestra la distribución de $n = 10$ elementos en $r = 4$ células).

Según se ve, el problema se reduce a la determinación del número de métodos mediante los cuales se pueda trazar $r - 1$ líneas en $n - 1$ intervalos entre los elementos: este número es igual a $\binom{n-1}{r-1}$. A este mismo tipo de problemas se refieren los siguientes: hallar el número de métodos para pintar, usando r colores, n objetos iguales (por ejemplo, bolas); hallar el número de r -combinaciones con repetición, en las cuales se emplea cada elemento.

2. Los elementos del conjunto N de distribuyen en las células de R de tal modo, que pueden haber células vacías. El método de resolución de los problemas de este tipo es, en lo principal, el mismo. Al conjunto de elementos N se le agregan r "elementos vacíos" simbólicos. En este caso el problema se reduce a la determinación del número de métodos de trazar $r - 1$ líneas en $n + r - 1$ intervalos entre los elementos. Este número será

$$\binom{n+r-1}{r-1} = \binom{n+r-1}{n}.$$

Entre los problemas de este tipo hay, por ejemplo, el siguiente: hallar el número de soluciones de la ecuación

$$x_1 + x_2 + \dots + x_r = n$$

en x_i : $i = 1, 2, \dots, r$ enteros no negativos.

3. En fin, a esta clase se refieren los problemas relacionados con el cálculo del número de muestras de un n -conjunto.

Los dos tipos restantes de problemas, (C) y (D), donde resultan no distinguibles las células para el llenado, representan, al tratar de resolverlos, dificultades mucho más considerables. Se llaman, habitualmente, por el nombre colectivo de particiones no ordenadas. Los problemas del tipo C, donde son indistinguibles sólo las células, mientras que los elementos a distribuir son distinguibles, permiten, por otra parte, su resolución. Indiquemos algunas fórmulas.

1. En aquellos casos en que no se admite, al realizar el llenado, células vacías y cuando se toma en consideración el orden en que los elementos caen en las células, el número buscado de distribuciones es igual a

$$A_n^r = \frac{n!}{r!} \binom{n-1}{r-1}. \quad (1)$$

2. Si la partición antecedente se modifica de tal modo que se admiten 1, 2, ..., r células vacías, entonces el número buscado de distribuciones es

$$(A')_n^r = A_n^r + A_n^{r-1} + \dots + A_n^1. \quad (2)$$

3. Si no hay células vacías, mientras que el orden de disposición de los elementos en las células no se toma en consideración, el número de distribuciones será

$$B_n^r = \frac{1}{r!} \sum_{\substack{s_1 + \dots + s_r = n \\ s_j \geq 1}} \frac{n!}{s_1! s_2! \dots s_r!} = S(n, r). \quad (3)$$

El número $S(n, r)$ se llama número de Stirling de 2º género (véase más abajo § 2.3).

4. Cuando en el caso 3 se admiten 1, 2, ..., r células vacías, el número de distribuciones es igual a

$$(B')_n^r = B_n^r + B_n^{r-1} + \dots + B_n^1. \quad (4)$$

Aduzcamos un esbozo de la demostración de las afirmaciones citadas. Efectivamente, supongamos que en el caso 1 se distribuyen elementos de un conjunto $N = \{\alpha_1, \alpha_2, \dots, \alpha_n\}$. Las distribuciones L tienen la forma $L = \{\alpha_1, \alpha_2, \dots, \alpha_{i_1} \mid \alpha_{i_1+1}, \alpha_{i_1+2}, \dots, \alpha_{i_2} \mid \dots \alpha_{i_{r-1}+1}, \dots, \alpha_{i_{r-1}+2}, \dots, \alpha_n\}$. El número $|L|$ de estas distribuciones se puede calcular mediante dos procedimientos:

a) $|L| = n! \binom{n-1}{r-1}$, es decir, el número de permutaciones en el conjunto N se multiplica por el número de distribuciones de $r-1$ líneas en $n-1$ intervalos;

b) $|L| = A_n^r \cdot r!$, es decir, el número buscado se multiplica por el número de reenumeraciones posibles de las células.

Al igualar entre sí a) y b), obtenemos la fórmula (1). Apliquemos en el caso 3, al igual que en el caso 1, dos métodos para calcular el número de distribuciones en diferentes células:

$$\begin{aligned} B_n^r \cdot r! &= \sum_{\substack{s_1 + \dots + s_r = n \\ s_i \geq 0}} \binom{n}{s_1} \binom{n-s_1}{s_2} \dots \binom{n-s_1-s_2-\dots-s_{r-1}}{s_r} = \\ &= \sum \frac{n!}{s_1!(n-s_1)!} \frac{(n-s_1)!}{s_2!(n-s_1-s_2)!} \dots \frac{(n-s_1-s_2-\dots-s_{r-1})!}{s_r!(n-s_1-s_2-\dots-s_r)!} = \\ &= \sum \frac{n!}{s_1! s_2! \dots s_r!}, \end{aligned}$$

de donde se deduce (3).

Las fórmulas (2) y (4) se desprenden, evidentemente, de las (1) y (3), respectivamente.

Estudiemos, por fin, el caso (D), es decir, la clase de problemas sobre las distribuciones, cuando los conjuntos N y R se componen ambos de elementos indistinguibles. Dichos problemas resultaron ser más difíciles y la teoría de su resolución aún no está elaborada completamente. La interpretación más conocida del caso dado la representa el problema teórico-numérico sobre la partición de los números naturales en sumandos naturales.

Notemos, ante todo, que los problemas de la clase (D) no pueden equipararse con los problemas de las clases anteriores, como se ha hecho más arriba: las conexiones resultan ser mucho más complejas y no se logra hallar una expresión analítica conveniente para obtener el número buscado.

Para calcular el número de particiones referentes a los problemas de tipo (D), sirve por ahora de medio principal el siguiente método recurrente.

Supongamos que un n -conjunto S se divide en k partes no vacías $a_1, a_2, \dots, a_k$, con la particularidad de que $|a_i| \geq 1, i = 1, 2, \dots, k$. Denotemos con $P_k(n)$ el número de tales particiones. Es obvio, que $P_1(n) = 1$; $P_k(k) = 1, P_k(n) = 0$ para $n < k$.

$$\sum_{i=1}^k a_i = n. \quad (5)$$

Admitamos que $|a_1| \geq |a_2| \geq \dots \geq |a_k|$ (renumeremos, si es necesario, las partes de la partición). Está claro que

$$\sum_{i=1}^k (a_i - 1) = n - k. \quad (6)$$

Se ha obtenido la partición del $(n - k)$ -conjunto en partes cuyo número es $\leq k$ (la igualdad tiene lugar, si cualquier a_i contiene no menos de 2 elementos). Según la regla de la suma, el número de tales particiones es igual a $\sum_{i=1}^k P_i(n - k)$, y, en virtud de la igualdad (6), dicho número es igual al número de particiones del n -conjunto en k partes, es decir, a

$$\sum_{i=1}^k P_i(n - k) = P_k(n). \quad (7)$$

Teniendo en cuenta los valores (5), esta fórmula recurrente permite obtener sucesivamente los valores para $P_k(n)$, reuniéndolos, si es necesario, en la tabla 1.1.

Para los valores pequeños de k podemos obtener fórmulas de $P_k(n)$, por ejemplo

$$P_1(n) = 1; P_2(2) = 1; P_2(1) = 0;$$

$$P_2(n) = P_2(n - 2) + P_1(n - 2) = P_2(n - 2) + 1;$$

de donde

$$P_2(n) = \frac{n}{2}, \text{ si } n \text{ es par};$$

$$P_2(n) = \frac{n-1}{2}, \text{ si } n \text{ es impar}.$$

Pero, ya para $k = 3$ las fórmulas se hacen bastante engorrosas. La investigación del comportamiento de las magnitudes $P_k(n)$ y $P(n) = \sum_{k=1}^n P_k(n)$ (el

Tabla 1.1

k	n	Números $P_k(n)$									
		1	2	3	4	5	6	7	8	9	10
1		1	1	1	1	1	1	1	1	1	1
2			1	1	2	2	3	3	4	4	5
3				1	1	2	3	4	5	7	8
4					1	1	2	3	5	6	9
5						1	1	2	3	5	7
6							1	1	2	3	5
7								1	1	2	3
8									1	1	2
9										1	1
10											1
$P(n) = \sum_k P_k(n)$		1	2	3	5	7	11	15	22	30	42

número de toda clase de particiones del número n) para grandes valores de k está ligada con dificultades considerables. Se ha hallado el valor aproximado

$$P_k(n) \sim \frac{1}{k!} \binom{n-1}{k-1},$$

el cual en la práctica resulta suficiente. Para $P(n)$ queda determinada una relación recurrente

$$P(n) = P(n-1) + P(n-2) - P(n-5) - P(n-7) + \dots \\ \dots + (-1)^{k-1} P\left(n - \frac{3k^2 \pm k}{2}\right) + \dots$$

y están calculados los primeros valores sucesivos de esta magnitud.

Todas estas cuestiones y otras semejantes se analizan en la teoría de las particiones (véase [16]).

Aproximadamente desde la mitad del siglo anterior, merced a los esfuerzos de Darfi, Ferrers, Sylvester, y más tarde, de MacMahon, en la teoría de las particiones entró la interpretación de las mismas con ayuda de los grafos puntuales. Por ejemplo, la partición $29 = 7 + 7 + 5 + 3 + 3 + 2 + 2$ está expuesta en la fig. 1.4.

Las partes de las particiones se disponen, como regla, de arriba abajo en el orden de decrecimiento. Del análisis de los correspondientes grafos n -puntuales (llamados también grafos de Ferrers) pueden directamente obtenerse los resultados siguientes.

1. El número de particiones de un n -conjunto en las que la mayor parte

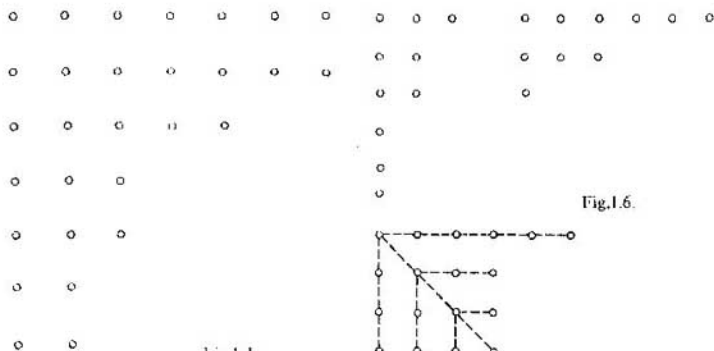


Fig. 1.4.

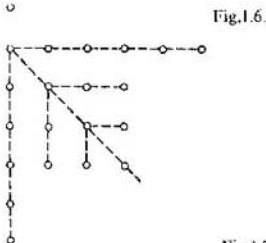


Fig. 1.6.

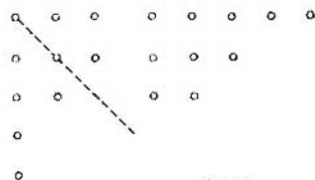


Fig. 1.5

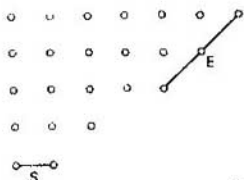


Fig. 1.7.

Fig. 1.8.

tiene k elementos, es igual al número de partes, es decir, a $P_k(n)$. Este teorema se demuestra por transposición del grafo de Ferrers respecto de la diagonal principal; tales grafos se denominan conjugados. Así, por ejemplo, en la fig. 1.5 el conjunto de 10 elementos está partido en 5 partes: $10 = 3 + 3 + 2 + 1 + 1$; al efectuar la transposición, obtenemos: $10 = 5 + 3 + 2$, es decir, la partición del 10-conjunto en la cual la parte mayor contiene 5 elementos (una situación análoga se muestra en la fig. 1.6: $10 = 3 + 2 + 2 + 1 + 1 + 1$ antes de la transposición, y $10 = 6 + 3 + 1$, después de la transposición).

2. El número de particiones autoconjugadas de un n -conjunto (una partición se llama autoconjugada, si el grafo de Ferrers que le corresponde es simétrico respecto de la diagonal principal) es igual al número de particiones del mismo conjunto en subconjuntos desiguales que se componen de un número impar de elementos. El teorema recíproco es también lícito.

Así, por ejemplo, a la partición autoconjugada del 20-conjunto $20 = 6 + 4 + 4 + 4 + 1 + 1$ (fig. 1.7) le corresponde biunívocamente la partición $20 = 11 + 5 + 3 + 1$, si los puntos de la primera columna dispuestos debajo de la diagonal los trasladamos a la primera fila, de la segunda columna, a la segunda fila, etc.

3. El número de particiones de un n -conjunto en partes diferentes es igual al número de particiones del mismo conjunto en partes, compuestas de un número impar de elementos.

Sea dada una partición del n -conjunto (por ejemplo, $n = 34$) en componentes impares ($34 = 5 + 5 + 5 + 5 + 3 + 3 + 3 + 1 + 1 + 1 + 1$). Todas las componentes impares iguales se reúnen en grupos (4 de cinco, 3 de tres y 5 unidades) y se escriben los números de sus repeticiones (4, 3 y 5) en el sistema binario ($4 = 2^2$; $5 = 2^2 + 2^0$; $3 = 2^1 + 2^0$). Escribamos la nueva partición, teniendo en cuenta la representación binaria obtenida ($34 = 5 \cdot 4 + 3 \cdot 3 + 1 \cdot 5 = 5 \cdot 2^2 + 3 \cdot 2^1 + 3 \cdot 2^0 + 1 \cdot 2^2 + 1 \cdot 2^2 + 1 \cdot 2^0 = 20 + 6 + 3 + 4 + 1$). Este procedimiento es siempre factible, pues cualquier número se escribe de un modo único en el sistema binario. Siempre se puede razonar también del modo inverso.

4. Si Q_n y Q'_n son los números de particiones del n -conjunto en un número par e impar, respectivamente, de las partes desiguales entre sí, entonces

$$Q_n = \begin{cases} Q'_n, & \text{si } n \neq \frac{k}{2} (3k \pm 1); \\ Q'_k + (-1)^k, & \text{si } n = \frac{n}{2} (3k \pm 1); \end{cases}$$

donde $k = 1, 2, \dots$

Supongamos que el grafo de Ferrers de la partición del n -conjunto en partes desiguales tiene la forma mostrada en la fig. 1.8. S es la parte menor de la partición y E , un conjunto de puntos (línea) dispuestos, a partir de la parte mayor, bajo un ángulo de 45° . Si $|S| \leq |E|$, entonces S se traslada al conjunto E ; en cambio, si $|S| > |E|$, viceversa, E se traslada a la parte menor S . Como resultado de tales traslaciones, tendremos el cambio de la paridad del número de partes desiguales de la partición. A toda partición par se le pone en correspondencia una impar y, además, biunívocamente: $Q_n = Q'_n$. Recomendamos que el lector mismo realice dicha operación con el siguiente ejemplo: $7 + 6 + 5 + 3 + 2 \leftrightarrow 8 + 7 + 5 + 3$.

Sin embargo, la operación no será posible, si las líneas S y E se intersecan y

$$|S| = |E|, \text{ o bien } |S| = |E| + 1.$$

Sea $|E| = k$. Entonces, en el primero de los casos exclusivos tendremos

$$n = k + (k + 1) + \dots + (2k - 1) = \frac{k}{2} (3k - 1);$$

y en el segundo:

$$n = (k + 1) + (k + 2) + \dots + 2k = \frac{k}{2} (3k + 1);$$

lo que demuestra nuestra afirmación.

1.5. SISTEMAS DE CONJUNTOS

Más arriba definimos la asignatura del análisis combinatorio, dimos una determinada información referente a la teoría de los conjuntos discretos finitos, indicamos las operaciones principales en las investigaciones combinatorias, explicamos los métodos de cálculo del número de entes combinatorios fundamentales, es decir, de muestras y ordenaciones, distribuciones y llenados. Todo lo dicho se realizó sólo para los conjuntos linealmente ordenados, y no para todos los casos posibles. Estas limitaciones fueron determinadas por el carácter introductor del capítulo.

No sería justo, sin embargo, concluir que los problemas tratados en el capítulo presente también son de importancia limitada y que el papel de los mismos es puramente pedagógico. La teoría general de los conjuntos discretos finitos es una asignatura matemática en desarrollo con sus propios problemas específicos. Además ocurre frecuentemente que a las investigaciones teóricas de los conjuntos finitos se reducen los problemas altamente prácticos.

Tomemos, por ejemplo, uno de los problemas sobre las particiones de un n -conjunto finito S en subconjuntos disjuntos. Analicemos las particiones π_1 y π_2 . Designemos, para un elemento $s \in S$ por $f_i(s)$ el número de elementos en el bloque de partición π_i , el cual contiene el elemento s ($i = 1, 2$). Llamemos conjugadas las particiones π_1 y π_2 , si los pares ordenados de números $(f_1(s), f_2(s))$ son distintos para todo $s \in S$.

Ejemplo. Sea $S = \{1, 2, 3, 4, 5, 6\}$. Definamos dos particiones:

$$\pi_1 : S = \{1, 2, 3\} \cup \{4, 5\} \cup \{6\};$$

$$\pi_2 : S = \{1, 4, 6\} \cup \{2, 5\} \cup \{3\}.$$

Obtendremos para ellas

i	1	2	3	4	5	6
$f_1(i), f_2(i)$	{3, 3}	{3, 2}	{3, 1}	{2, 3}	{2, 2}	{1, 3}

Por cuanto todos los pares ordenados $\{f_1(i), f_2(i)\}$ son distintos, π_1 y π_2 serán particiones conjugadas del conjunto S . Con relación a estas particiones conjugadas se ha demostrado [35] que un par de estas particiones existe, si y sólo si $n \neq 2, 5, 9$. Este resultado se ha obtenido al resolver el siguiente problema sobre la identificación de los cables telefónicos multifilares.

Sea dado un cable de n conductores indistinguibles. Se pide fijar en sus extremos A y B los bornes $A_i, B_i; i = 1, 2, \dots, n$, que correspondan a cada conductor. El método de resolución consiste en unir los grupos de extremos de los conductores por un lado del cable y en probar el paso de la corriente que se mide por otro lado del cable. Explicaremos esto más detalladamente con un ejemplo de un cable de seis conductores.

Tomemos el extremo A con bornes $A_1, A_2, \dots, A_6$, y unamos sus bornes de acuerdo con la partición π_1 (posición I en la fig. 1.9). Supongamos que

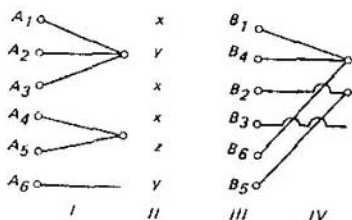


Fig. 1.9.

como resultado de la comprobación por el lado del extremo B (con bornes $B_1, B_2, \dots, B_6$) se ha obtenido la situación que se indica en la fig. 1.9 mediante la posición II. Los bornes indistinguibles B_1, B_2, B_3 se denotan con x , los bornes B_4 y B_5 con y , y el B_6 con z ; unamos los bornes en el extremo B de acuerdo con la partición π_2 (posición III). A continuación, al desconectar los bornes en el extremo A , comprobamos cuántos y cuáles conductores por el lado A quedan conectados en el extremo B . Admitamos, por ejemplo, que el conductor $A_1 B_1$ resultó conectado con un solo borne en el extremo B . Por cuanto sabemos que el borne A_1 pertenece al conjunto $\{A_1, A_2, A_3\}$, mientras que el borne en el extremo B integra el conjunto de dos bornes unidos (a saber, B_2 y B_3), y por cuanto el par $\{f_1(i), f_2(i)\} = \{3, 2\}$ corresponde solamente a $i = 2$, concluimos que A_1 está conectado con B_2 . Análogamente, si el borne A_5 está conectado con dos otros bornes por el lado B (es decir, el conductor $A_5 B_5$ pertenece al conjunto de tres conductores que quedaban conectados), entonces A_5 debe ser conectado con B_4 (dado que al par $(2, 3)$ corresponde $i = 4$), etc.

Este algoritmo fue extendido también al caso general. Sea I un n -conjunto de números enteros $I = \{1, 2, \dots, n\}$. Supongamos que para I existen dos particiones conjugadas:

$$\pi_1 : I = P_1 \cup P_2 \cup \dots \cup P_k;$$

$$\pi_2 : I = P'_1 \cup P'_2 \cup \dots \cup P'_k.$$

En el extremo A unamos primeramente los conductores de acuerdo con la partición π_1 , es decir, conductores en los cuales los números de los bornes pertenecen a cierto $P_l \in \pi_1$; $l = 1, 2, \dots, k$. En el extremo B elegimos por comprobación los subconjuntos $S_1, S_2, \dots, S_k$ de bornes de aquellos conductores que fueron conectados en el extremo A , y enumeremos los bornes en el extremo B de un modo tal que, cualquiera que sea S_r , el conjunto de índices de los bornes B_i que se encuentran en S_r sea exactamente un subconjunto de π_1 . Ahora unamos los conductores en el extremo B en grupos $T_1, T_2, \dots, T_k$, donde T_j consta de todos aquellos conductores con bornes B_i , para los cuales $i \in P'_j$. Separando todas las conexiones en el extremo A , escogemos los conductores por el lado de A que están conectados en el extremo B .

Realizada esta operación, podemos encontrar los bornes "izquierdo" y "derecho" para un mismo conductor. En efecto, sea A_u algún conductor

que hemos cogido, y supongamos que en el extremo B dicho conductor integra el grupo T_j , el cual tiene, digamos, p elementos (de donde hallamos este número p). Por cuanto se sabe el número de elementos en el subconjunto S_i , en el que está contenido A_u (admitamos que es igual a q), entonces, basándonos en el procedimiento de construir S_i y T_j , y en que todos los pares $(p, q) = \{f_1(u), f_2(u)\}$ son diferentes, podemos encontrar el único B_{iu} tal que A_u y B_{iu} sean extremos de un mismo conductor.

En muchas ramas de las matemáticas y en sus aplicaciones se plantean y se resuelven problemas que no sólo se reducen al análisis de los conjuntos discretos y de sus sistemas, sino que, además se enuncian en términos de los mismos. Así, por ejemplo, sucede en la teoría de los autómatas finitos, en la técnica de cómputo discreta, problemas aplicados del álgebra, etc. Describamos unos cuantos problemas tipo que se encuentran muy a menudo.

Familias de Sperner. Se dice que los subconjuntos $S_1, S_2, \dots, S_m$ de un conjunto finito S forman una familia de Sperner, siempre que ninguno de ellos está contenido en el otro. Sea $|S| = n$. ¿Cuál será el número máximo posible m de términos en la familia de Sperner? La respuesta a esta pregunta lleva el nombre del teorema de Sperner (véase cap. 8): $m = \binom{n}{[n/2]}$ ¹⁾.

Sistemas separadores. Este concepto fue introducido por A. Rényi [36] al analizar los problemas de la teoría de información. Un sistema de subconjuntos $\{S_1, S_2, \dots, S_m\}$ de un conjunto finito S se denomina separador, si en el mismo, para cualesquiera dos elementos distintos del conjunto S , existe un subconjunto S_i que contiene sólo un elemento de todos los mencionados. A. Rényi planteó el problema de hallar un sistema separador mínimo bajo la condición de que cada subconjunto de este sistema consta exactamente de un número dado de elementos. Este problema fue analizado en [37].

Problemas sobre los subconjuntos que se intersecan. Existen varios problemas en los que se introducen limitaciones en la potencia de los propios subconjuntos $S_1, S_2, \dots, S_m \subseteq S$ y de sus intersecciones. Se requiere determinar el número máximo (m) de subconjuntos que satisfagan dichas condiciones para $n = |S|$ fijo. Con este motivo recomendamos la obra [38]. Demos a conocer el resultado clásico de Erdős, Chao Ko, Rado [39]. Supongamos que

- 1) cada uno de los subconjuntos $S_1, S_2, \dots, S_m$ contiene no más de k elementos, donde $k \leq n/2$;
- 2) ninguno de los subconjuntos está contenido en el otro;
- 3) cualesquiera dos subconjuntos se intersecan. En este caso el número

¹⁾ Con $[x]$ se denota el número máximo entero que no sobrepasa x ; con $\{x\}$, el número entero mínimo, superior o igual a x .

máximo posible de subconjuntos será $\binom{n-1}{k-1}$. Precisamente este número de términos tiene el sistema de todos los k -subconjuntos del n -conjunto S en los que está contenido cierto elemento fijo $s \in S$.

Recubrimientos y empaques. El problema de construcción de los mejores códigos conduce al siguiente problema combinatorio: hallar el número máximo m , para el cual existe un sistema de subconjuntos $\{S_1, S_2, \dots, S_m\}$ de r elementos del n -conjunto S , donde $|S_i \cap S_j| < t$ para cualesquiera $1 \leq i < j \leq m$. Dicho de otro modo, se requiere que cada t -subconjunto del conjunto S se contenga a lo sumo en uno de los subconjuntos del sistema. Este problema se denomina *problema de empaques*. También se analiza con frecuencia el problema inverso. Exijamos que cada t -subconjunto se contenga no menos que en uno de los subconjuntos del sistema; se pregunta qué cantidad mínima de r -subconjuntos del n -conjunto S es necesaria para que se pueda formar tal sistema. Este problema se denomina *problema de recubrimientos*. Al igual que el problema de empaques, él está resuelto por ahora sólo en algunos casos particulares (por ejemplo, para $r = 3, t = 2$; $r = 4, t = 2$). Si en lugar de los subconjuntos $S_1, S_2, \dots, S_m$ estudiamos complementos de los mismos $S'_1, S'_2, \dots, S'_m$, donde $S'_i = S \setminus S_i$, y ponemos $k = n - r, t = n - r$, obtendremos otra forma del problema de recubrimientos: ¿qué número mínimo de t -subconjuntos del n -conjunto S es necesario para que en cualquier k -subconjunto de conjunto S se contenga por lo menos uno de los t -subconjuntos elegidos. Este número lleva el nombre de *número de Turán* $T(n, k, t)$. En el año 1941 Turán [40] demostró que

$$T(n, k, 2) = mn - \frac{m(m+1)}{2} (k-1) \text{ para } m \leq \frac{n}{k-1} \leq m+1.$$

Pese a que la formulación es sencilla, el problema de hallar los números de Turán resulta ser, en el caso general, exclusivamente difícil. Para $t \geq 3$ se han obtenido pocos resultados. Se conoce (véase [41]) que

$$T(n, k, t) = n - (k-1) \text{ para } 1 \leq \frac{n}{k-1} \leq \frac{t}{t-1}.$$

$$T(n, k, t) = \left\lfloor \frac{\left(\left\lceil \frac{3t}{2} \right\rceil - 1 \right)n - \left\lceil \frac{3t}{2} \right\rceil (k-1)}{\left\lceil \frac{t}{2} \right\rceil} \right\rfloor$$

$$\text{para } \frac{t}{t-1} \leq \frac{n}{k-1} \leq \left\lceil \frac{\left\lceil \frac{3t}{2} \right\rceil}{\left\lceil \frac{t}{2} \right\rceil - 2} \right\rceil.$$

La propiedad $\mathcal{B}$. Suele decirse que un sistema de subconjuntos $S_1, S_2,$

..., $S_m \subseteq S$ posee la propiedad $\mathcal{B}$, si existe tal partición del conjunto $S = S' \cup S''$ ($S' \cap S'' = \emptyset$), que $S_i \not\subseteq S'$, $S_i \not\subseteq S''$ ($i = 1, 2, \dots, m$). Se imponen las restricciones $|S| = n$, y $|S_1| = |S_2| = \dots = |S_m| = k$, y se busca el número mínimo $m = m(n, k)$ de subconjuntos en un sistema que no posea la propiedad $\mathcal{B}$. Este problema se analiza en el cap. 7.

Una clase importante más de problemas sobre los sistemas de conjuntos está relacionada con el teorema de Ramsey al que se dedica un párrafo aparte del capítulo 3.

El estudio de los sistemas de conjuntos es el problema principal en el análisis combinatorio. De los éxitos de dicho estudio depende tanto el enriquecimiento de la teoría de la combinatoria, como también la ampliación de los campos de aplicación.

CAPÍTULO 2 FUNCIONES GENERATRICES

Durante mucho tiempo el contenido del análisis combinatorio lo constituía el cálculo del número de configuraciones de determinados tipos. Una parte de la teoría combinatoria que estudia estos problemas sigue jugando hoy día un papel importante en las aplicaciones.

En §§ 1.3 y 1.4 hemos considerado los métodos directos ("elementales") de cálculo. El presente capítulo se dedica a los métodos indirectos, con ayuda de los cuales se calcula la cantidad de configuraciones combinatorias.

2.1. FUNDAMENTOS DEL MÉTODO DE FUNCIONES GENERATRICES

El método de las funciones generatrices (generadoras) es uno de los más desarrollados en el análisis combinatorio. Las ideas fundamentales de este método fueron enunciadas por primera vez al final del siglo XVIII en las obras de Laplace referentes a la teoría de las probabilidades. Expliquemos las mismas con el siguiente ejemplo sencillo. Veamos el producto del número finito de binomios lineales

$$\prod_{r=1}^n (1 + x_r t) = \sum_{r=0}^n |a_r| t^r, \quad (1)$$

donde

$$a_r(x_1, \dots, x_n) = \sum_{1 \leq i_1 < i_2 < \dots < i_r \leq n} x_{i_1} x_{i_2} \dots x_{i_r}$$

no son nada más que funciones simétricas elementales de las variables $x_1, \dots, x_n$. Observemos que a los sumandos del coeficiente a_r se les puede asignar las r -combinaciones de n elementos $x_1, \dots, x_n$. La expresión (1) se llamará *numerador* de las r -combinaciones de n elementos. Si ponemos en (1) $x_i = 1$ para $i = 1, \dots, n$, obtendremos

$$\sum_{r=0}^n \binom{n}{r} t^r = (1 + t)^n, \quad (2)$$

puesto que $a_r(1, \dots, 1)$ es el número de r -combinaciones de n elementos. Desarrollemos la función $(1 + t)^n$ en potencias de t según la fórmula de Taylor:

$$(1 + t)^n = \sum_{r=0}^n \frac{n!}{r!(n-r)!} t^r. \quad (3)$$

(Este mismo resultado puede ser demostrado también por inducción respecto de n .)

De (2) y (3) se deduce

$$\sum_{r=0}^n \binom{n}{r} t^r = \sum_{r=0}^n \frac{n!}{r!(n-r)!} t^r. \quad (4)$$

Al igualar entre sí los coeficientes de iguales potencias de t en (4), obtenemos de nuevo (esta vez, analíticamente) el resultado del capítulo 1: el número de r -combinaciones de n elementos es igual a $\frac{n!}{r!(n-r)!}$. Al aplicar este resultado del capítulo 1 a la fórmula (2), obtendremos una demostración más de la identidad (3). Semejante método de determinar los coeficientes del desarrollo de las funciones era generalmente aceptado en la primera mitad del siglo XIX y se llamaba análisis combinatorio, a diferencia del análisis matemático al cual se referían los métodos analíticos de obtención de los desarrollos.

En la expresión (2) la función $f(t) = (1+t)^n$ está biunívocamente relacionada con una sucesión de los números

$$\left\{ \binom{n}{r} \right\}, \quad r = 0, 1, 2, \dots, n.$$

Tal relación resulta muy útil: atribuyendo en la fórmula (2) diferentes valores particulares a la variable t , se pueden obtener muchas identidades importantes. Así, por ejemplo, para $t = 1$ y $t = -1$ tenemos

$$2^n = \sum_{r=0}^n \binom{n}{r} = 1 + n + \binom{n}{2} + \binom{n}{3} + \dots + \binom{n}{n},$$

$$0 = \sum_{r=0}^n (-1)^r \binom{n}{r} = 1 - n + \binom{n}{2} - \binom{n}{3} + \dots + (-1)^n \binom{n}{n};$$

respectivamente. La adición y sustracción de estas expresiones término a término nos da

$$\sum_{r=0}^{\lfloor \frac{n}{2} \rfloor} \binom{n}{2r} = \sum_{r=0}^{\lfloor \frac{n-1}{2} \rfloor} \binom{n}{2r+1} = 2^{n-1},$$

mientras que la simple división de los factores

$$(1+t)^n = (1+t)^m (1+t)^{n-m}$$

conduce a la identidad

$$\binom{n}{r} = \sum_{k=0}^r \binom{n-m}{k} \binom{m}{r-k},$$

que se conoce como convolución de Vandermonde. Por fin, al sustituir en (2) $t = a/b$, y al multiplicar por b^n los miembros primero y segundo, obtenemos seguidamente el *teorema binomial* (la fórmula binomial):

$$(a + b)^n = \sum_{r=0}^n \binom{n}{r} a^r b^{n-r}.$$

debido a lo cual los números $\binom{n}{r}$ se llaman *coeficientes binomiales*.

Ejercicio. Demuéstrese, por inducción respecto a m , el teorema binomial:

$$\left(\sum_{i=1}^m a_i \right)^n = \sum \frac{n!}{n_1! n_2! \dots n_m!} a_1^{n_1} a_2^{n_2} \dots a_m^{n_m},$$

donde la suma se calcula respecto a todas las soluciones de la ecuación $n_1 + n_2 + \dots + n_m = n$ en números enteros no negativos; los coeficientes en el segundo miembro se llaman polinomiales y ya aparecieron en el capítulo I como números de las $(n_1, n_2, \dots, n_m)$ -particiones del n -conjunto.

La función $f(t) = (1 + t)^n$ se denomina *función generatriz* de una sucesión de números $\left\{ \binom{n}{r} \right\}$, o bien, más brevemente, función generadora del número de r -combinaciones de n elementos, $r = 0, 1, 2, \dots, n$.

Examinemos ahora una sucesión numérica $a = (a_0, a_1, a_2, \dots)$, o bien, de otro modo, la función a_n de un argumento de número entero n . A esta función le corresponde biunívocamente la serie

$$f_a(t) = \sum_{n=0}^{\infty} a_n t^n, \quad (5)$$

la cual es más cómoda y simple en las operaciones, particularmente cuando ella converge a una función que posee una forma analítica conveniente. La serie $f_a(t)$ se llama *función generatriz* de la sucesión a . Así pues, las funciones generatrices permiten pasar del análisis de magnitudes aisladas (por ejemplo, de las r -combinaciones para un valor particular r) al análisis de sus sucesiones e, incluso, de las clases de sucesiones.

Para la mayoría de los problemas combinatorios la serie (5) es finita. Si esta serie es, sin embargo, infinita y el radio de su círculo de convergencia es igual a cero, las operaciones con ella son posibles sólo dentro del álgebra de las series de potencias formales, el cual se analizará en el párrafo que viene.

2.2. TIPOS DE FUNCIONES GENERATRICES Y DE NUMERADORES

Sea R cierto anillo con la unidad. El anillo $S(R)$ de sucesiones sobre R y el anillo $R[[t]]$, isomorfo a $S(R)$, de series de potencias formales sobre R se definen del modo siguiente. Los elementos del anillo $S(R)$ son las sucesiones

$$\{a\} = \{(a_0, a_1, a_2, \dots)\}, \quad (1)$$

y los elementos correspondientes del anillo $R[[t]]$ son las series

$$\{F_a(t)\}, F_a(t) = \sum_{r=0}^{\infty} a_r t^r. \quad (2)$$

Se llama suma de las sucesiones $a = (a_0, a_1, \dots)$ y $b = (b_0, b_1, \dots)$ a la sucesión

$$c = a + b = (a_0 + b_0, a_1 + b_1, \dots) = (c_0, c_1, \dots),$$

y se llama suma de las series $F_a(t)$ y $F_b(t)$, pertenecientes a la clase (2), a la serie $F_c(t) = F_a(t) + F_b(t)$:

$$F_c(t) = \sum_{r=0}^{\infty} c_r t^r,$$

donde $c_r = a_r + b_r$.

Se denomina producto (o convolución) de las sucesiones a y b de la clase (1) a la sucesión $a \times b = d = (d_0, d_1, \dots)$, en la cual

$$d_r = a_0 b_r + a_1 b_{r-1} + \dots + a_r b_0, \quad r = 0, 1, \dots, \quad (3)$$

y producto (convolución) de las series $F_a(t)$ y $F_b(t)$ de la clase (2), a la serie

$$F_d(t) = F_a(t) \times F_b(t) = \sum_{r=0}^{\infty} d_r t^r,$$

donde d_r se determina según la fórmula (3).

Definamos ahora: el cero en la clase (1) como una sucesión (nula)

$$0 = (0, 0, \dots),$$

entonces el cero en la clase (2) es la serie correspondiente a O

$$F_0(t) = 0;$$

la unidad en la clase (1) como una sucesión (unitaria)

$$e = (1, 0, 0, \dots),$$

entonces la unidad en la clase (2) será una serie correspondiente a e :

$$F_e(t) = 1.$$

Por fin, un elemento inverso para $a \in S(R)$ respecto de la adición en la clase (1) es $-a = (-a_0, -a_1, \dots)$, y el elemento correspondiente inverso pa-

ra $F_a(t)$ en la clase (2) será

$$-F_a(t) = F_{-a}(t) = \sum_{r=0}^{\infty} (-a_r)t^r.$$

Es fácil ver que todos los axiomas del anillo para $S(R)$ y $R[[t]]$ tienen lugar.

Sea a_0 un elemento invertible del anillo R ; buscamos $a^{-1} = a'$, partiendo de la condición $a \times a' = e$, es decir,

$$\begin{aligned} a_0 a'_0 &= 1, \\ a_1 a'_0 + a_0 a'_1 &= 0, \\ &\dots \dots \dots \\ a_k a'_0 + a_{k-1} a'_1 + a_{k-2} a'_2 + \dots + a_0 a'_k &= 0, \\ &\dots \dots \dots \end{aligned}$$

de donde encontramos a'_k ($k = 0, 1, \dots$) de las primeras $k+1$ ecuaciones, empleando el cálculo sucesivo por el método de Gauss (o bien por el de Cramer). Por consiguiente, sólo para las sucesiones a , tales que a_0 es invertible, existen a^{-1} y $F_a^{-1}(t)$ en los anillos $S(R)$ y $R[[t]]$.

En los anillos $S(R)$ y $R[[t]]$ puede introducirse la diferenciación D : para $a = (a_0, a_1, \dots)$, $Da = (a_1, 2a_2, \dots, na_n, \dots)$,

$$DF_a(t) = \sum_{n=1}^{\infty} na_n t^{n-1},$$

y la integración $\int$:

$$\int a = \left(0, a_0, \frac{a_1}{2}, \dots, \frac{a_n}{n+1}\right), \quad \int F_a(t) = \sum_{n=0}^{\infty} \frac{a_n}{n+1} t^{n+1}.$$

Ejercicio 1. Demuéstranse las siguientes propiedades de las aplicaciones D y $\int$ en $S(R)$:

- 1) $D(a+b) = Da + Db$; 2) $\int(a+b) = \int a + \int b$;
- 3) $D\int a = a$; 4) $\int Da = a$;
- 5) $D(a \times b) = (Da) \times b + a \times Db$;

y las propiedades correspondientes en el anillo $R[[t]]$.

Si el anillo R es un álgebra sobre el campo P , entonces, al introducir en $S(R)$ y en $R[[t]]$ las operaciones de multiplicación por $\alpha \in P$:

$$\alpha a = (\alpha a_0, \alpha a_1, \dots), \quad \alpha F_a(t) = \sum_{r=0}^{\infty} (\alpha a_r)t^r,$$

convertimos $S(R)$ y $R[[t]]$ en álgebras isomórfas.

Volvamos ahora a las sucesiones que aparecen en los problemas combinatorios, es decir, al álgebra $S(\mathbf{R})$ de sucesiones sobre el campo $\mathbf{R}$ de números reales. Para la sucesión $a \in S(\mathbf{R})$, la serie $F_a(t) \in R[[t]]$ lleva el nombre de *función generatriz ordinaria*, mientras que la propia álgebra $R[[t]]$ se deno-

mina álgebra de Cauchy. Las funciones generatrices ordinarias se usan al analizar las familias de sucesiones cuyos elementos están constituidos por las funciones de las r -muestras no ordenadas (r -combinaciones).

Para la construcción de los mismos objetos combinatorios (de las r -combinaciones) emplearemos el álgebra $R(x; t) = R[x_1, x_2, \dots][[t]]$ de series de potencias formales sobre el álgebra $R[x_1, x_2, \dots]$ de polinomios de las variables $x = \{x_1, x_2, \dots\}$ sobre el campo de números reales. El álgebra $R(x; t)$ se llamará *numeradora*, y las series, cuyos coeficientes serán polinomios que enumeran los objetos combinatorios construidos por nosotros, *numeradores* de dichos objetos.

Recordemos que la aplicación $M: A_1 \rightarrow A_2$, realizada del álgebra A_1 sobre el campo P en el álgebra A_2 sobre P , se denomina *operador lineal*, si $M(a + b) = Ma + Mb$ y $M(\alpha a) = \alpha Ma$ para cualesquiera $a, b \in A_1, \alpha \in P$. La aplicación M se denomina *operador multiplicativo*, si $M(ab) = (Ma)(Mb)$ para cualesquiera $a, b \in A_1$. Usaremos con frecuencia la aplicación $T: R(x; t) \rightarrow R[[t]]$ que convierte todos los x_i en 1.

Ejercicio 2. Demuéstrese que la aplicación T es un operador multiplicativo lineal.

Ejercicio 3. Demuéstrese que las aplicaciones D y $\int$ sobre el álgebra $R(x; t)$ son operadores lineales.

EjemPlo 1. Hállense el numerador y la función generatriz para las r -combinaciones de n elementos.

Este problema ya lo hemos resuelto en el § 2.1: el numerador de las r -combinaciones de n elementos está dado por la serie (1) en el álgebra numeradora $R(x; t)$, y la función generatriz, por la serie (4) en la subálgebra $R[[t]]$ de la misma.

EjemPlo 2. Hállense el numerador y la función generatriz para las r -combinaciones con repeticiones del tipo $(\Lambda_1, \dots, \Lambda_n)$ de n elementos, donde Λ_i es un subconjunto del conjunto N_0 de números enteros no negativos; el elemento a_i puede estar presente en las r -combinaciones λ veces, siempre que $\lambda \in \Lambda_i$.

Al elemento a_i del n -conjunto S pongámosle en correspondencia una variable x_i en $R(x; t)$, entonces, a la aparición del elemento a_i en la r -combinación corresponderá, o bien λ_1 veces, ..., o bien λ_2 veces, la serie $x_i^{\lambda_1} t^{\lambda_1} + x_i^{\lambda_2} t^{\lambda_2} + \dots$. Por eso, el numerador buscado tendrá la expresión

$$\sum_{r=0}^{\infty} a_r t^r = \prod_{i=1}^n \sum_{\lambda \in \Lambda_i} x_i^{\lambda} t^{\lambda}. \quad (4)$$

La función generatriz tiene en este caso la forma

$$F(t) = \prod_{i=1}^n \sum_{\lambda \in \Lambda_i} t^{\lambda} \quad (5)$$

(a los miembros primero y segundo de (4) se ha aplicado el operador lineal T , que convierte todos los x_i en 1). Al desarrollar (5) en potencias de t , llegamos a que el número de r -combinaciones del tipo $(\lambda_1, \dots, \lambda_n)$ de n elementos

es igual al número de soluciones de la ecuación $y_1 + \dots + y_n = r$ con las incógnitas $y_i \in \Lambda_i$, $i = 1, \dots, n$.

EJEMPLO 3. Hállense el numerador y la función generatriz, para las r -combinaciones con repeticiones ilimitadas de n elementos.

Reduciremos nuestro problema al anterior, si ponemos $\Lambda_i = N_0$, $i = 1, \dots, n$. El numerador (4) adopta la forma:

$$\sum_{r=0}^{\infty} a_r t^r = \prod_{i=1}^n (1 + x_i t + x_i^2 t^2 + \dots) = \sum_{i=1}^n \frac{1}{1 - x_i t}.$$

De aquí (ó de (5)) encontramos la función generatriz:

$$\begin{aligned} f(t) &= (1 + t + t^2 + \dots)^n = (1 - t)^{-n} = \\ &= \sum_{r=0}^{\infty} \frac{(-n)(-n-1)\dots(-n-r+1)}{n!} (-1)^r t^r = \\ &= \sum_{r=0}^{\infty} \binom{n+r-1}{r} t^r = \sum_{r=0}^{\infty} \binom{-n}{r} (-t)^r, \end{aligned}$$

donde se ha puesto por definición

$$\binom{-n}{r} = (-1)^r \binom{n+r-1}{r}.$$

De aquí tenemos la sucesión $b = (b_0, b_1, \dots)$, donde $b_r = \binom{n+r-1}{r}$ es el número de r -combinaciones con repetición de n elementos. Este resultado concuerda con el obtenido antes (véase § 1.3).

EJEMPLO 4. Hállense el numerador y la función generatriz de las r -combinaciones con repetición, en las cuales figura por lo menos un elemento de cada tipo.

Pongamos, en las condiciones del ejemplo 2, $\Lambda_i = N = \{1, 2, \dots\}$, $i = 1, \dots, n$. Entonces, de (4) obtenemos el numerador buscado:

$$\sum_{r=0}^{\infty} a_r t^r = \prod_{i=1}^n (x_i t + x_i^2 t^2 + \dots) = t^n \prod_{i=1}^n \frac{x_i}{1 - x_i t}.$$

De aquí tenemos la función generatriz:

$$f(t) = (t + t^2 + \dots)^n = t^n (1 - t)^{-n} = t^n \sum_{r=0}^{\infty} \binom{n+r-1}{r} t^r.$$

Al realizar la sustitución $n + r = k$, obtenemos

$$f(t) = \sum_{k=n}^{\infty} \binom{k-1}{k-n} t^k = \sum_{k=n}^{\infty} \binom{k-1}{n-1} t^k.$$

Por consiguiente, el número de r -combinaciones buscadas es igual a 0 para $r < n$, y a $\binom{r-1}{n-1}$, para $r \geq n$.

EJEMPLO 5. Hállense el numerador y la función generatriz para las r -combinaciones de n elementos, donde se admite sólo un número par de apariciones para cada uno de los elementos.

Suponiendo en las condiciones del ejemplo 2, $\Lambda_i = \{0, 2, 4, \dots\}$, $i = 1, \dots, n$, tenemos de (4) el numerador buscado:

$$\sum_{r=0}^{\infty} a_r t^r = \prod_{i=1}^n (1 + x_i^2 t^2 + x_i^4 t^4 + \dots) = \prod_{i=1}^n \frac{1}{1 - x_i^2 t^2}.$$

De aquí obtenemos la función generatriz

$$f(t) = (1 + t^2 + t^4 + \dots)^n = (1 - t^2)^{-n} = \sum_{r=0}^{\infty} \binom{n+r-1}{r} t^{2r}.$$

Ahora podemos obtener una identidad interesante: por cuanto $(1 - t^2)^{-n} = (1 - t)^{-n}(1 + t)^{-n}$, la comparación de los coeficientes de las potencias de t en los miembros primero y segundo nos da:

$$\sum_{k=0}^r (-1)^k \binom{n+r-k-1}{r-k} \binom{n+k-1}{k} = \begin{cases} 0, & \text{para } r \text{ impar,} \\ \binom{n+s-1}{s}, & \text{para } r = 2s. \end{cases}$$

Ejercicio 4. Hállense el numerador y la función generatriz para las r -combinaciones de n elementos, donde se admiten a lo sumo j repeticiones de cada elemento.

Pasemos ahora a la construcción de la teoría analítica para la enumeración de las r -muestras ordenadas o de las r -permutaciones.

En los márgenes del álgebra $\mathbf{R}[x_1, x_2, \dots][[t]]$ de series de potencias formales con los coeficientes del álgebra de polinomios $\mathbf{R}[x_1, x_2, \dots]$ sobre un campo $\mathbf{R}$ de variables $x_1, x_2, \dots$, que no conmutan entre sí, examinemos la expresión

$$\frac{1}{n!} \sum_{\pi \in S_n} \prod_{i=1}^n (1 + x_{\pi(i)} t) = \sum_{r=0}^n a_r^*(x_1, x_2, \dots) \frac{t^r}{r!}, \quad (6)$$

en la que la suma en el primer miembro se toma por todas las sustituciones de grupo simétrico S_n de sustituciones de un n -conjunto. No es difícil ver que en el coeficiente de t^r del desarrollo en potencias de t de la expresión

$$\sum_{\pi \in S_n} \prod_{i=1}^n (1 + x_{\pi(i)} t)$$

el número de apariciones del monomio $x_{i_1} \dots x_{i_r}$ es igual al número de sustituciones de $\pi \in S_n$ tales, que $\pi^{-1}(i_1) < \pi^{-1}(i_2) < \dots < \pi^{-1}(i_r)$. Las sustituciones de esta índole pueden construirse del modo siguiente: elegimos r lugares para las preimágenes de los elementos $i_1, \dots, i_r$ empleando $\binom{n}{r}$ méto-

dos, y los demás $n - r$ elementos pueden permutarse mediante $(n - 1)!$ métodos. Por eso, el monomio que se examina aparece sólo $\binom{n}{r} (n - r)! = n!/r!$ veces. Por consiguiente, el polinomio $a_r(x_1, x_2, \dots)$ en (6) se compone de monomios, correspondientes a todas las r -permutaciones de elementos $x_1, \dots, x_n$. La serie (6) se denominará numerador exponencial de las r -permutaciones de n elementos.

Ahora, suponiendo en (6) $x_i = 1, i = 1, \dots, n$ (aplicamos a los términos primero y segundo el operador multiplicativo T que convierte todos los x_i en 1), obtenemos

$$(1 + t)^n = \sum_{r=0}^n P(n, r) \frac{t^r}{r!}, \quad (7)$$

donde $P(n, r)$ es el número de r -permutaciones de n elementos. La función $f(t) = (1 + t)^n$ se llamará función generatriz exponencial para el número de r -permutaciones de n elementos. Observemos que la fórmula (7) podía ser deducida inmediatamente de la fórmula (2) § 2.1, puesto que $P(n, r) = a_r(1, \dots, 1)$. Al desarrollar $(1 + t)^n$ en potencias de t , obtenemos

$$\sum_{r=0}^n P(n, r) \frac{t^r}{r!} = \sum_{r=0}^n \frac{n!}{(n - r)!} \frac{t^r}{r!},$$

y de aquí hallamos otra vez el número de r -permutaciones de n elementos: $P(n, r) = n!/(n - r)!$.

Sea R un anillo con la unidad. El anillo $S^*(R)$ de sucesiones exponenciales sobre R y el anillo $R^*[[t]]$ (isomorfo a $S^*(R)$) de series de potencias formales sobre R se construyen del modo siguiente. El anillo $S^*(R)$, siendo un grupo aditivo, coincide con el grupo aditivo del anillo $S(R)$. La operación de multiplicación en $S^*(R)$ será la convolución binomial de las sucesiones: $ab = d = (d_0, d_1, \dots)$, donde $a = (a_0, a_1, \dots)$, $b = (b_0, b_1, \dots)$,

$$d_r = \sum_{i=0}^r \binom{r}{i} a_i b_{r-i}. \quad (8)$$

La unidad y el cero en $S^*(R)$ son los mismos que en el anillo $S(R)$. El elemento inverso respecto de la multiplicación para la sucesión $a = (a_0, a_1, \dots)$ existe sólo en el caso de invertibilidad del elemento a_0 , y se calcula a partir del sistema de ecuaciones, análogo al sistema, ya analizado, de ecuaciones en el anillo $S(R)$.

Ejercicio 5. Demuéstrese que en $S^*(R)$ están cumplidos todos los axiomas del anillo, y la aplicación $\sigma = (a_0, a_1, \dots, a_n, \dots) \rightarrow \sigma' = (a_0, a_1, \dots, a_n/n!, \dots)$ es un isomorfismo de los anillos $S(R)$ y $S^*(R)$.

Como elementos del anillo $R^*[[t]]$ intervienen las series de potencias exponenciales

$$[E_a(t)], E_a(t) = \sum_{r=0}^{\infty} a_r \frac{t^r}{r!},$$

correspondientes a las sucesiones $a = (a_0, a_1, \dots)$. Las operaciones de adición y multiplicación se determinan del modo siguiente:

$$E_a(t) + E_b(t) = E_c(t) = \sum_{r=0}^{\infty} c_r \frac{t^r}{r!},$$

$$b = (b_0, b_1, \dots), c_r = a_r + b_r;$$

$$E_a(t) \times E_b(t) = E_d(t) = \sum_{r=0}^{\infty} d_r \frac{t^r}{r!},$$

donde d_r se halla por la fórmula (8). El cero y la unidad en $R^*[[t]]$ son los mismos que en el anillo $R[[t]]$; el elemento inverso respecto de la adición se determina análogamente; el elemento inverso respecto de la multiplicación se calcula igual que en el anillo $S^*(R)$. Si R es el álgebra sobre el campo P , entonces (lo mismo que en el caso de $S(R)$) el anillo $S^*(R)$ se transforma en el álgebra sobre P . Una construcción análoga se realiza también para el anillo $R^*[[t]]$ sobre el álgebra R .

Ejercicio 6. Realícense para el anillo $R^*[[t]]$ las construcciones y demostraciones omitidas en el texto.

De la definición del anillo $R^*[[t]]$ se ve que en el caso de las series convergentes el producto formal coincide con el ordinario. La serie $E_a(t)$ recibe el nombre de *función generatriz exponencial* para la sucesión $a \in S^*(R)$, y la propia álgebra $R^*[[t]]$ se denomina álgebra de funciones generatrices exponenciales.

Ejercicio 7. Muéstrese que las álgebras $\bar{R}^*[[t]]$ y $R^*[[t]]$ son isomorfas.

De nuevo con el fin de realizar la construcción algebraica de las r -muestras ordenadas, examinemos el álgebra $\bar{R}^*(x; t) = (R[x_1, x_2, \dots])^*[[t]]$ de series de potencias exponenciales sobre el álgebra $\bar{R}[x_1, x_2, \dots]$ de polinomios con variables no conmutables, la cual llamaremos álgebra numeradora exponencial, y el álgebra $R^*(x; t) = (R[x_1, x_2, \dots])^*[[t]]$ de series de potencias exponenciales sobre el álgebra $R[x_1, x_2, \dots]$ de polinomios habituales que llamaremos álgebra numeradora exponencial reducida. Denotemos con T^* la aplicación de $\bar{R}^*(x; t)$ en $R^*[[t]]$ que reemplaza todos los x_i por 1.

Ejercicio 8. Demuéstrese que T^* es un operador lineal multiplicativo.

Designemos con U la aplicación que traslada una serie de $\bar{R}^*(x; t)$ en la misma serie de $R^*(x; t)$, es decir, que permite que las variables x_i conmuten entre sí.

Ejercicio 9. Demuéstrese que U es un operador multiplicativo lineal.

Designemos con T^* la aplicación $R^*(x; t)$ en $R^*[[t]]$ que reemplaza todos los x_i por 1.

Ejercicio 10. Demuéstrese que T^* es un operador multiplicativo lineal.

Veamos la aplicación $V: R^*(x; t) \rightarrow R^*(x; t)$, la cual todo monomio $x_{i_1} x_{i_2} \dots x_{i_r}$ con la especificación $(\lambda_1, \dots, \lambda_n)$ (esto quiere decir que x_i se encuentra λ_i veces, x_2 , λ_2 veces, etc.) en el coeficiente de $t^k/k!$ lo transforma en el polinomio

$$\frac{\lambda_1! \dots \lambda_n!}{r!} \sum x_{j_1} \dots x_{j_r}$$

donde la suma se toma respecto de todas las permutaciones $(j_1, \dots, j_r)$ de los elementos $i_1, \dots, i_r$.

Ejercicio 11. Demuéstrese que V es un operador lineal, mas la propiedad de multiplicatividad para él no tiene lugar.

Ejercicio 12. Demuéstrese que UV es un operador idéaleo sobre el álgebra $R^*(x; t)$.

A toda r -permutación $i_1 i_2 \dots i_r$ de elementos $1, \dots, n$ con repeticiones pongamos en correspondencia en el álgebra numeradora exponencial un monomio $x_{i_1} \dots x_{i_r} t^r / r!$. La serie correspondiente $p(x_1, x_2, \dots; t)$ para la clase dada de permutaciones se llamará *numerador exponencial* de dicha clase. La imagen $U_p(x_1, x_2, \dots; t)$, al actuar el operador U , la llamaremos *numerador exponencial reducido* de la clase dada de permutaciones, y la serie $\tilde{T}^* p(x_1, x_2, \dots; t)$, *función generatriz exponencial*.

EjemPlo 6. Hállense el numerador exponencial, el numerador exponencial reducido y la función generatriz exponencial para una r -permutación de n elementos.

El numerador exponencial ya lo hemos encontrado en la forma (6); el numerador exponencial reducido es el polinomio

$$\prod_{i=1}^n (1 + x_i t),$$

y la función generatriz exponencial para el número de r -permutaciones de n elementos la hemos hallado en la forma de (7).

EjemPlo 7. Hállense el numerador exponencial, el numerador exponencial reducido y la función generatriz exponencial para las r -permutaciones con repeticiones del tipo $(\lambda_1, \dots, \lambda_n)$ de n elementos.

Por $(\lambda_1, \dots, \lambda_n)$ -permanente (véase [42]) de una $(r \times n)$ -matriz

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{r1} & \dots & a_{rn} \end{pmatrix}$$

con los elementos del anillo R se entenderá un elemento

$$\text{Per}_{(\lambda_1, \dots, \lambda_n)} A = \sum a_{1i_1} a_{2i_2} \dots a_{ri_{r_1}}$$

donde la suma se toma respecto de todas las r -permutaciones $(i_1, \dots, i_r)$ de elementos $1, \dots, n$ con la repetición del tipo $(\lambda_1, \dots, \lambda_n)$, o bien, dicho de otro modo, en el producto pueden figurar λ_i elementos de la columna i sólo en el caso en que $\lambda_i \in A_i$ si la suma tiene un conjunto vacío de sumandos,

entonces pongámosla igual a cero. Ahora podemos escribir el numerador exponencial buscado en la forma

$$p(x_1, x_2, \dots; t) = \sum_{r=0}^{\infty} \text{Per}(\Lambda_1, \dots, \Lambda_n) X_{r,n} \frac{t^r}{r!},$$

donde $X_{r,n}$ es una $(r \times n)$ -matriz con r filas iguales $(x_1, \dots, x_n)$. El numerador exponencial reducido $Up(x_1, x_2, \dots; t)$ puede representarse ahora del modo siguiente:

$$Up(x_1, x_2, \dots; t) = \sum_{r=0}^{\infty} \sum_{\substack{r_1 + \dots + r_n = r \\ r_1 \in \Lambda_1, \dots, r_n \in \Lambda_n}} \frac{r!}{r_1! \dots r_n!} x_1^{r_1} \dots x_n^{r_n} \frac{t^r}{r!}, \quad (9)$$

de donde

$$Up(x_1, x_2, \dots; t) = \prod_{i=1}^n \left(\sum_{\lambda \in \Lambda_i} x_i^\lambda \frac{t^\lambda}{\lambda!} \right). \quad (10)$$

Por fin, la función generatriz exponencial tiene la forma

$$\tilde{T}_p^*(x_1, x_2, \dots; t) = T^* Up(x_1, x_2, \dots; t) = \prod_{i=1}^n \left(\sum_{\lambda \in \Lambda_i} \frac{t^\lambda}{\lambda!} \right),$$

por lo cual

$$p(1, 1, \dots; t) = \sum_{r=0}^{\infty} \sum_{\substack{r_1 + \dots + r_n = r \\ r_1 \in \Lambda_1, \dots, r_n \in \Lambda_n}} \frac{r!}{r_1! \dots r_n!} \frac{t^r}{r!}.$$

De aquí llegamos a que el número de r -permutaciones de n elementos con repeticiones del tipo $(\Lambda_1, \dots, \Lambda_n)$ es igual a

$$\sum_{\substack{r_1 + \dots + r_n = r \\ r_1 \in \Lambda_1, \dots, r_n \in \Lambda_n}} \frac{r!}{r_1! \dots r_n!}.$$

Podemos hallar también el numerador exponencial partiendo del numerador exponencial reducido (10), al representarlo en la forma canónica (9) y al aplicar el operador lineal V .

EJEMPLO 8. Hállense el numerador exponencial, el numerador exponencial reducido y la función generatriz exponencial para las r -permutaciones con repeticiones ilimitadas de n elementos.

En este caso, en las condiciones del ejemplo antecedente $\Lambda_i = N_0$ para $i = 1, \dots, n$, y

$$\text{Per}(N_0, \dots, N_0) X_{r,n} = (x_1 + \dots + x_n)^r.$$

Por eso el numerador exponencial tiene la forma de un exponente

$$p(x_1, x_2, \dots; t) = \sum_{r=0}^{\infty} (x_1 + \dots + x_n)^r \frac{t^r}{r!} = \exp((x_1 + \dots + x_n)t).$$

La misma forma tiene en este caso el numerador exponencial reducido, mientras que la función generatriz exponencial se obtiene en la siguiente forma

$$p(1, 1, \dots; t) = \exp(nt) = \sum_{r=0}^{\infty} n^r \frac{t^r}{r!}.$$

Así pues, el número de r -permutaciones con repeticiones ilimitadas de n elementos es igual a n^r , lo que concuerda con el resultado del § 1.3. La forma de la función generatriz de este ejemplo predetermina precisamente el término "exponencial" que se usa en la teoría analítica de enumeraciones de las r -permutaciones.

EjemPlo 9. Hállense el numerador exponencial, el numerador exponencial reducido y la función generatriz exponencial para las r -permutaciones de n elementos con repeticiones, donde cada elemento ha de aparecer por lo menos una vez.

Para este caso, en las condiciones del ejemplo 7 tenemos $\Lambda_i = \mathbf{N}$, $i = 1, \dots, n$, por consiguiente, el numerador exponencial reducido tendrá la forma

$$p^*(x_1, x_2, \dots; t) = \prod_{i=1}^n \left(\sum_{\lambda=1}^{\infty} \frac{(x_i t)^{\lambda}}{\lambda!} \right) = \prod_{i=1}^n (\exp(x_i t) - 1).$$

Ahora podemos escribir la función exponencial generatriz:

$$\begin{aligned} p^*(1, 1, \dots; t) &= (e^t - 1)^n = \sum_{k=0}^n (-1)^k \binom{n}{k} e^{t(n-k)} \\ &= \sum_{r=0}^{\infty} \frac{t^r}{r!} \sum_{k=0}^n (-1)^k \binom{n}{k} (n-k)^r, \end{aligned}$$

los números correspondientes, por lo tanto, son iguales a

$$p(n, r) = \sum_{k=0}^n (-1)^k \binom{n}{k} (n-k)^r. \quad (11)$$

Por fin, el numerador exponencial puede ser determinado con ayuda del operador lineal V :

$$p(x_1, x_2, \dots; t) = V p^*(x_1, x_2, \dots; t) = V \prod_{i=1}^n (e^{x_i t} - 1).$$

Ejercicio 13. Hállense el numerador exponencial, el numerador exponencial reducido y la función generatriz exponencial para las r -permutaciones con repeticiones de n elementos, donde cada elemento puede aparecer sólo un número par de veces.

La función generatriz de Dirichlet para una sucesión de números $a = (a_1, a_2, \dots)$ es una serie formal

$$D_a(t) = \sum_{r=1}^{\infty} \frac{a_r}{r^t}.$$

Determinemos, igual que lo hicimos anteriormente, las operaciones de adición y multiplicación para estas series. Se llamará suma de $D_a(t)$ con $D_b(t)$ a la serie

$$D_a(t) + D_b(t) = D_c(t) = \sum_{r=1}^{\infty} \frac{c_r}{r^t}, \quad c_r = (b_1, b_2, \dots), \quad \dots$$

donde $c_r = a_r + b_r$, y producto de las series mencionadas, a la serie siguiente:

$$D_a(t) \times D_b(t) = D_d(t) = \sum_{r=1}^{\infty} \frac{d_r}{r^t},$$

donde

$$d_r = \sum_{ij=r} a_i b_j = \sum_{i|r} a_i b_{r/i}$$

(la adición se realiza según los divisores enteros r). De aquí se ve con facilidad que $O = (0, 0, \dots)$ define el cero con relación a la adición; $e = (1, 0, 0, \dots)$ define la unidad con relación a la multiplicación: $D_e(t) = 1$; el elemento inverso respecto de la multiplicación $D_a^{-1}(t)$ se busca a partir de la igualdad

$$D_a(t) \times D_a^{-1}(t) = D_e(t) = 1,$$

o bien, que es lo mismo, del sistema de ecuaciones:

$$\begin{aligned} a_1 a'_1 &= 1, \\ a_2 a'_1 + a_1 a'_2 &= 0, \\ a_3 a'_1 + a_2 a'_2 + a_1 a'_3 &= 0, \\ a_4 a'_1 + a_2 a'_2 + a_1 a'_4 &= 0, \\ &\dots \end{aligned}$$

de donde tenemos, en particular: $a'_1 = \frac{1}{a_1}$, $a'_2 = -\frac{a_2}{a_1^2}$, $a'_3 = -\frac{a_3}{a_1^2}$, $a'_4 = \frac{a_2^2}{a_1^3} - \frac{a_4}{a_1^2}$, etc., es decir, es necesaria y suficiente la invertibilidad del elemento a_1 . No es difícil comprobar que la totalidad de funciones generatrices de Dirichlet es también un anillo. Si definimos la multiplicación natural de la serie $D_a(t)$ por un número real, dicho anillo será el álgebra $\mathbf{R}\{t\}$

sobre el campo $\mathbf{R}$. El álgebra de Dirichlet $D(\mathbf{R})$ de sucesiones, isomorfa a la primera, se determina de un modo análogo.

La noción de función generatriz de Dirichlet surgió en la teoría de los números, donde es de amplio uso la zeta función de Riemann:

$$\xi(t) = \sum_{r=1}^{\infty} \frac{1}{r^t}.$$

Las tres formas de las funciones generatrices introducidas más arriba son fundamentales para el análisis combinatorio, pero están lejos de ser únicas. La diversidad de las formas que tienen funciones generatrices, se debe a los diversos planteamientos de los problemas. Veamos, como ejemplo, un problema sobre la construcción de las funciones generatrices para alojamientos y llenados, a saber, el problema del tipo (A_2) (véase § 1.4).

Se tienen un n -conjunto de diferentes elementos y un m -conjunto de diferentes células de especificación $(n_1, n_2, \dots, n_m)$. Se requiere construir la función generatriz para el número de alojamientos de n elementos en m células. Supongamos que el símbolo

$$x_1^{n_1} x_2^{n_2} \dots x_m^{n_m}$$

significa que en la i -ésima célula se alojan n_i elementos, $i = 1, \dots, m$.

Si $n = 1$, es decir, si se tiene sólo un elemento, entonces el correspondiente símbolo de alojamiento es x_i , siempre que el elemento está colocado en la i -ésima célula, $i = 1, \dots, m$. De conformidad con la regla de la suma, la posibilidad de alojar un elemento en cualquiera de m células se describe por el polinomio

$$x_1 + x_2 + \dots + x_m.$$

Si se tienen n elementos, de los cuales cada uno puede ser colocado en cualquiera de m células, entonces el polinomio correspondiente será

$$(x_1 + x_2 + \dots + x_m)^n,$$

mientras que la correspondiente función generatriz se representará por una función exponencial

$$\begin{aligned} E(t) &= \sum_{n=0}^{\infty} (x_1 + x_2 + \dots + x_m)^n \frac{t^n}{n!} = \exp \left(t \sum_{i=1}^m x_i \right) = \\ &= \exp \left(\sum_{i=1}^m t x_i \right) = \prod_{i=1}^m \exp(t x_i). \end{aligned}$$

Si se trata sólo del número de alojamientos posibles, suponemos $x_i = 1$, $i = 1, 2, \dots, m$, y, en este caso, la función generatriz adopta la forma

$$E(t) = \exp(mt).$$

De acuerdo con el teorema polinomial tenemos:

$$\left(\sum_{i=1}^m x_i\right)^n = \sum \frac{n!}{n_1! n_2! \dots n_m!} x_1^{n_1} x_2^{n_2} \dots x_m^{n_m},$$

de donde obtenemos que el número de alojamientos de n diferentes elementos en m diferentes células con una $(n_1, n_2, \dots, n_m)$ -especificación será igual a

$$\frac{n!}{n_1! n_2! \dots n_m!}.$$

2.3. APARATO OPERACIONAL DEL MÉTODO DE FUNCIONES GENERATRICES

El uso de las funciones generatrices, permite enfocar con un grado suficiente de generalidad problemas combinatorios del tipo enumerativo. Sin embargo, las funciones generatrices, construidas para diferentes tipos de muestras resultaron ser, en la mayor parte bastante engorrosas. Por esta razón, para una notación más cómoda de las funciones generatrices se emplean operadores especiales, cálculos simbólicos y números y funciones especiales.

Sea una sucesión $a = (a_0, a_1, \dots)$, la que se escribirá de otro modo: $a = a(n)$, es decir, como función de un argumento de número entero $n (n = 0, 1, 2, \dots)$. Para el conjunto de sucesiones $\{a(n)\}$ se emplean con mayor frecuencia los siguientes operadores especiales:

a) el operador de desplazamiento E :

$$Ea(n) = a(n+1), \quad n = 0, 1, 2, \dots, \\ E^{-1}a(n) = a(n-1), \quad n = 1, 2, \dots, \quad E^{-1}a(0) = 0;$$

b) el operador de diferencia Δ :

$$\Delta a(n) = a(n+1) - a(n);$$

c) el operador mediador δ :

$$\delta a(n) = \frac{a(n-1) + a(n+1)}{2}, \quad n = 1, 2, \dots, \quad \delta a(0) = \frac{a(1)}{2}.$$

Estos operadores están unidos por medio de ciertas relaciones, por ejemplo:

$$\Delta a(n) = (E - I)a(n), \\ \delta a(n) = \frac{1}{2} (E + E^{-1})a(n),$$

donde I es un operador idéntico.

Las reiteradas aplicaciones de los operadores a los elementos de la sucesión pueden conducir a ciertas fórmulas útiles y cómodas. Por ejemplo, denotemos con n^r una sucesión con término general $a(n) = n^r$. La notación

Ln^r , en la que L es un operador lineal en el conjunto de sucesiones se entenderá como $(La)(n)$, donde $a(n) = n^r$. Entonces,

$$E^n(0) = n; E^n(0^r) = n^r; \Delta 0^r = 1^r - 0^r = 1^r; \\ \Delta^2 0^r = 2^r - 2; \Delta^3 0^r = 3^r - 3 \cdot 2^r + 3,$$

etc. Los números $\Delta^n 0^r$ llevan el nombre de *De Morgan*.

Como un ejemplo de empleo de los operadores en las fórmulas combinatorias aduzcamos el problema sobre el número de r -permutaciones con repeticiones de un n -conjunto a condición de que cada elemento aparezca no menos de una vez.

Reduzcamos el número $p(n, r)$, obtenido en el § 2.2, a la forma (11) del modo siguiente:

$$p(n, r) = \sum_{i=0}^n \binom{n}{i} (-1)^i (n-i)^r = \sum_{i=0}^n \binom{n}{i} (-1)^i E^{n-i} 0^r = (E-1)^n 0^r = \\ = \Delta^n 0^r.$$

Demostremos que

$$\Delta^n 0^{r+1} = n \Delta^n 0^r + n \Delta^{n-1} 0^r.$$

En efecto,

$$\Delta^n 0^{r+1} = \sum_{i=0}^n \binom{n}{i} (-1)^i (n-i)^{r+1} = n \sum_{i=0}^n \binom{n}{i} (-1)^i (n-i)^r - \\ - \sum_{i=0}^n \binom{n}{i} (-1)^i (n-i)^r i.$$

El primer sumando no es otra cosa que $n \Delta^n 0^r$, lo que se deduce con toda evidencia de lo expuesto anteriormente. En lo que se refiere al segundo sumando, está claro que al realizar la sustitución $i = j+1$, obtenemos

$$\sum_{i=0}^n \binom{n}{i} (-1)^i (n-i)^r i = - \sum_{j=0}^{n-1} \binom{n}{j+1} (-1)^j (n-j-1)^r (j+1) \\ = - \sum_{j=0}^{n-1} \frac{n(n-1)! (j+1)}{(j+1)! (n-1-j)!} (-1)^j (n-j-1)^r = \\ = -n \sum_{j=0}^{n-1} \binom{n-1}{j} (-1)^j (n-1-j)^r = -n \Delta^{n-1} 0^r.$$

Ejercicio 1. Demuéstrese que el número de r -permutaciones de n elementos con repeticiones pares (incluido 0) ilimitadas es igual a $\delta^n 0^r$.

Las repeticiones frecuentes de las mismas expresiones condujeron a la aparición de números y funciones especiales. Los números especiales, por

ejemplo, se descubren y vuelven a descubrir en los diferentes apartados de las matemáticas: en el cálculo de las diferencias finitas, en la teoría de los números, en la teoría de las probabilidades y en la estadística matemática. He aquí algunos datos sobre dichos números.

Analicemos la factorial inferior $(t)_n = t(t-1)\dots(t-n+1)$ como una función generatriz:

$$(t)_n = \sum_{k=0}^n s(n, k) t^k, \quad n > 0.$$

Los coeficientes $s(n, k)$ en este desarrollo se denominan *números de Stirling de primer género*. Por cuanto $(t)_k$ es un polinomio de grado k , entonces, podemos desarrollar t^n según el sistema de polinomios $(t)_0, (t)_1, \dots, (t)_n$:

$$t^n = \sum_{k=0}^n S(n, k) (t)_k, \quad n > 0. \quad (1)$$

Los coeficientes $S(n, k)$ se denominan *números de Stirling de segundo género*. Realicemos una definición adicional para ambos géneros de números:

$$(t)_0 = t^0 = s(0, 0) = S(0, 0) = 1.$$

Los números de Stirling se encuentran en muchos problemas. Volvamos, por ejemplo, al problema examinado más arriba sobre el número de n -permutaciones con repeticiones de k elementos a condición de que cada elemento figure en las muestras mencionadas por lo menos una vez (véase el ejemplo 9 del § 2.2). Resulta que la función generatriz en este caso tiene la forma

$$\begin{aligned} (e^t - 1)^k &= \sum_{i=0}^k \binom{k}{i} (-1)^i e^{t(k-i)} = \sum_{n=0}^{\infty} \frac{t^n}{n!} \sum_{i=0}^k \binom{k}{i} (-1)^i (k-i)^n = \\ &= \sum_{n=0}^{\infty} \Delta^k 0^n \frac{t^n}{n!}. \end{aligned}$$

Recordemos que este problema admite su interpretación como un problema sobre el número de alojamientos de n diferentes elementos en k células distinguibles tales que no queden células vacías. Para cualesquiera $t \geq n$ naturales tenemos

$$\begin{aligned} t^n &= E^t 0^n = (1 + \Delta)^t 0^n = \sum_{k=0}^t \binom{t}{k} \Delta^k 0^n = \sum_{k=0}^t (t)_k \frac{\Delta^k 0^n}{k!} = \\ &= \sum_{k=0}^n (t)_k \frac{\Delta^k 0^n}{k!}, \end{aligned}$$

puesto que $\Delta^k 0^n = 0$ para $k > n$ ($\Delta^k 0^n$ es el número de n -permutaciones de k elementos con repeticiones, donde cada elemento figura no menos de 1 vez). Por cuanto el polinomio de grado n cuenta con n raíces, entonces la

igualdad

$$t^n = \sum_{k=0}^n (t)_k \frac{\Delta^k 0^n}{k!}$$

será válida para todos los valores reales de t . Al tener presente el desarrollo (1), obtenemos

$$\Delta^k 0^n = k! S(n, k).$$

Esto quiere decir que el número de alojamientos de n elementos diferentes en k células distinguibles, cuando ninguna célula queda vacía (problema del tipo (A)), es igual a $k! S(n, k)$, y el correspondiente número para las células indistinguibles (problema del tipo (C)) (o lo que es lo mismo, el número de particiones de un n -conjunto en k partes no vacías) es igual a $S(n, k)$.

Los números de Stirling se encuentran a menudo en los razonamientos combinatorios, por lo cual daremos a conocer algunas observaciones más.

Se pueden obtener recurrencias útiles para dichos números, por ejemplo:

a) de $(t)_{n+1} = (t-n)(t)_n$ se deduce

$$s(n+1, k) = s(n, k-1) - ns(n, k);$$

$$\begin{aligned} \text{b) } t^{n+1} &= \sum_{k=0}^{n+1} S(n+1, k)(t)_k = t \sum_{k=0}^n S(n, k)(t)_k = \\ &= \sum_{k=0}^n S(n, k)((t)_{k+1} + k(t)_k), \end{aligned}$$

de donde obtenemos

$$S(n+1, k) = S(n, k-1) + kS(n, k).$$

Con una de las dos expresiones

$$(t)_n = \sum_{k=0}^n s(n, k)t^k, \quad t^n = \sum_{k=0}^n S(n, k)(t)_k$$

sustituimos la otra:

$$t^n = \sum_{k=0}^n S(n, k) \sum_{m=0}^k s(k, m)t^m = \sum_{k=0}^n \sum_{m=0}^k S(n, k)s(k, m)t^m.$$

Al igualar los coeficientes de las mismas potencias de t , obtenemos

$$\sum_{k=m}^n S(n, k)s(k, m) = \delta_{n,m} = \begin{cases} 1, & \text{si } n = m, \\ 0, & \text{si } n \neq m. \end{cases}$$

El número $\delta_{n,m}$ se denomina *delta de Kronecker*.

He aquí una información acerca de los otros números especiales.

Se llaman *de Fibonacci* los números $f(n)$, donde

$$f(0) = f(1) = 1, f(n) = f(n-1) + f(n-2), n \geq 2.$$

Su función generatriz tiene la forma:

$$F(t) = 1 + t + 2t^2 + 3t^3 + 5t^4 + 8t^5 + 13t^6 + \dots$$

Si se toma en consideración la relación recurrente, llegamos a la siguiente expresión analítica para la función generatriz:

$$\begin{aligned} F(t) &= 1 + t + (f(1) + f(0))t^2 + \dots + (f(n-1) + f(n-2))t^n + \dots = \\ &= 1 + (t + t^2)F(t), \end{aligned}$$

de donde

$$F(t) = (1 - t - t^2)^{-1}.$$

Esta función puede ser representada también en la forma

$$F(t) = ((1 - at)(1 - bt))^{-1}, \quad (2)$$

donde a y b se hallan de las correlaciones

$$a + b = 1, ab = -1,$$

de donde

$$a = \frac{1 + \sqrt{5}}{2}, b = \frac{1 - \sqrt{5}}{2}.$$

Al desarrollar (2) en fracciones simples, obtenemos

$$F(t) = \frac{\alpha}{1 - at} + \frac{\beta}{1 - bt}, \quad \alpha = \frac{1 + \sqrt{5}}{2\sqrt{5}}, \quad \beta = -\frac{1 - \sqrt{5}}{2\sqrt{5}}$$

(α y β se hallan por el método de coeficientes indeterminados). De aquí obtenemos la fórmula de Binet

$$f(n) = \alpha a^n + \beta b^n = \frac{\left(\frac{1 + \sqrt{5}}{2}\right)^{n+1} - \left(\frac{1 - \sqrt{5}}{2}\right)^{n+1}}{\sqrt{5}}.$$

Los números de Fibonacci forman una sucesión específica que se encuentra a menudo en los problemas combinatorios. Por ejemplo, el número de Fibonacci $f(n)$, $n \geq 1$, es el número de $(n-1)$ -permutaciones de ceros y unidades que no contienen dos ceros seguidos.

De perfeccionamiento ulterior del aparato de funciones generatrices sirve el empleo del cálculo simbólico de Blissar. Demos a conocer la versión rígida del mismo que se basa en las ideas de Mullin y Rota [43] referentes a la utilización del aparato de operadores lineales con el objeto de fundamentar el cálculo de Blissar.

Recordemos que un operador lineal que actúa de un espacio vectorial en el campo sobre el cual está construido dicho espacio se denomina *funcional lineal*.

La idea fundamental del cálculo de Blissar consiste en la operación "de alzamiento de los subíndices": la función generatriz exponencial

$$E_a(t) = \sum_{n=0}^{\infty} a_n \frac{t^n}{n!}$$

de una sucesión $a = (a_0, a_1, \dots)$ se sustituye por la serie

$$\exp(at) = \sum_{n=0}^{\infty} a^n \frac{t^n}{n!}$$

y después de realizar las transformaciones indispensables con la participación de esta serie, las potencias se sustituyen de nuevo por los subíndices (en lugar de a^n escribimos a_n). Con el fin de fundamentar estas transformaciones introduzcamos el álgebra $\mathbf{R}^*(a; t) = (\mathbf{R}[a])^*[[t]]$ de series exponenciales formales con los coeficientes en el álgebra de polinomios $\mathbf{R}[a]$ de la variable a , y la funcional lineal $L_a: \mathbf{R}[a] \rightarrow \mathbf{R}$ tal que $L_a a^n = a_n$, prolongada, naturalmente, hasta el operador lineal $L_a: \mathbf{R}^*(a; t) \rightarrow \mathbf{R}^*[[t]]$, de suerte que $L_a \exp(at) = E_a(t)$.

Ejem p l o 1. Veamos una sucesión de los números de Bernoulli $B = (B_0, B_1, \dots)$, definida por su función exponencial

$$\sum_{n=0}^{\infty} B_n \frac{t^n}{n!} = \frac{t}{e^t - 1}.$$

Entonces, para la funcional lineal $L_B: \mathbf{R}[B] \rightarrow \mathbf{R}$ tal que $L_B B^n = B_n$, tenemos

$$L_B \exp(Bt) = \frac{t}{e^t - 1}. \quad (3)$$

De aquí tenemos

$$\begin{aligned} t &= L_B(\exp((B+1)t) - \exp(Bt)) = L_B \left(\sum_{n=0}^{\infty} ((B+1)^n - B^n) \frac{t^n}{n!} \right) = \\ &= L_B \left(\sum_{n=0}^{\infty} \left(\sum_{r=0}^{n-1} \binom{n}{r} B^r \right) \frac{t^n}{n!} \right) = \sum_{n=0}^{\infty} \left(\sum_{r=0}^{n-1} \binom{n}{r} B_r \right) \frac{t^n}{n!}, \end{aligned}$$

y obtenemos, de este modo, la fórmula recurrente

$$\sum_{r=0}^{n-1} \binom{n}{r} B_r = \begin{cases} 1 & \text{para } n = 1 \\ 0 & \text{para } n = 0, 2, 3, 4, \dots \end{cases}$$

Sustituyendo en (3) t por $-t$, obtenemos

$$L_B \exp(-Bt) = \frac{-t}{e^{-t} - 1} = \frac{te^t}{e^t - 1} = L_B \exp((B+1)t).$$

Iguando los coeficientes de las potencias de x , tenemos

$$L_B(B+1)^n = L_B(-B)^n,$$

es decir,

$$\sum_{r=0}^n \binom{n}{r} B_r = (-1)^n B_n. \quad (4)$$

De (3) y (4) obtenemos, ahora, que $B_n = (-1)^n B_n$ para $n \neq 1$, y $B_1 = -B_1 - 1$. Por consiguiente, $B_1 = -1/2$, $B_3 = B_5 = B_7 = \dots = 0$.

Ejercicio 2. Demuéstrese que en el álgebra $R^*(a; t)$ tienen lugar las propiedades de permutabilidad de los operadores lineales D , $\int$ y L_a :

1. $D L_a = L_a D$;
2. $\int L_a = L_a \int$.

Analícemos ahora el caso cuando en las expresiones que se consideran participan dos sucesiones $a = (a_0, a_1, \dots)$ y $b = (b_0, b_1, \dots)$. Entonces, para cada una de ellas construimos su funcional lineal: $L_a: a^n \rightarrow a_n$, y $L_b: b^n \rightarrow b_n$ (trabajamos ya en el álgebra $R^*(a, b; t) = (R[a, b])^*[[t]]$ con polinomios de dos variables que actúan en calidad de coeficientes). Veamos, por ejemplo, la sucesión $c = (c_0, c_1, \dots)$ que representa su convolución binomial. Suponiendo que $L_a(p(b)a^n) = p(b)L_a a^n$ para cualquier polinomio $p(b)$ (y, análogamente, para L_b), tenemos

$$c_n = \sum_{k=0}^n \binom{n}{k} a_k b_{n-k} = L_a L_b \sum_{k=0}^n \binom{n}{k} a^k b^{n-k} = L_a L_b (a+b)^n.$$

Introduciendo la funcional lineal $L_c: c^n \rightarrow c_n$, obtenemos de aquí

$$L_c c^n = L_a L_b (a+b)^n.$$

Para el producto de funciones generatrices exponenciales correspondientes tenemos

$$\begin{aligned} L_a \exp(ct) &= E_a(t) = E_a(t)E_b(t) = \\ &= L_a L_b \exp(at) \exp(bt) = L_a L_b \exp((a+b)t). \end{aligned}$$

La dificultad principal del cálculo clásico de Blissar consistía en que, al considerar la convolución

$$\sum_{k=0}^n \binom{n}{k} a_k a_{n-k}$$

después de alzar los subíndices:

$$\sum_{k=0}^n \binom{n}{k} a^k a^{n-k}$$

era necesario prohibir la multiplicación $a^k a^{n-k} = a^n$, puesto que dicha operación conduciría a un resultado absurdo:

$$\sum_{k=0}^n \binom{n}{k} a_k a_{n-k} = 2^n a_n.$$

Siguiendo las ideas de Guinand [44], esta posibilidad la eliminamos aquí exigiendo que, al alzar los subíndices de las convoluciones de las sucesiones iguales, se introduzcan para ellas diferentes variables. En el caso dado tomamos $L_a : a^n \rightarrow a_n$ y $L_{a'} : (a')^n \rightarrow a_n$. Entonces

$$\sum_{k=0}^n \binom{n}{k} a_k a_{n-k} = L_a L_{a'} (a + a')^n.$$

EJEMPLO 2. Demostremos la identidad de Euler para la suma de los productos de los números de Bernoulli B_n :

$$\sum_{r=1}^{n-1} \binom{2n}{2r} B_{2r} B_{2n-2r} = -(2n+1) B_{2n}.$$

Esta identidad puede reescribirse en la forma

$$\sum_{s=0}^m (-1)^s \binom{m}{s} B_s B_{m-s} = (1-m) B_m \quad (5)$$

(cuando m es impar, los miembros primero y segundo son nulos). Introducimos las funcionales lineales $L_B : B^n \rightarrow B_n$ y $L_{B'} : B'^n \rightarrow B_n$, entonces (5) puede escribirse en la siguiente forma:

$$L_B L_{B'} (B - B')^m = (1-m) L_B B^m. \quad (6)$$

Analicemos la función generatriz exponencial para el primer miembro de (6):

$$\begin{aligned} L_B L_{B'} \sum_{m=0}^{\infty} (B - B')^m \frac{t^m}{m!} &= L_B L_{B'} \exp((B - B')t) = \\ &= L_B L_{B'} \exp(Bt) \exp(-B't) = \frac{t}{e^t - 1} \frac{-t}{e^{-t} - 1} = \frac{t^2 e^t}{(e^t - 1)^2} = \\ &= -t^2 \frac{d}{dt} \left(\frac{L_B \exp(Bt)}{t} \right) = L_B (1 - Bt) \exp(Bt) = \\ &= L_B \left(\sum_{m=0}^{\infty} B^m \frac{t^m}{m!} - \sum_{m=0}^{\infty} B^{m+1} \frac{t^{m+1}}{m!} \right) = L_B \left(\sum_{m=0}^{\infty} (1-m) B^m \frac{t^m}{m!} \right). \end{aligned}$$

De aquí obtenemos la identidad (6), lo que se trataba de demostrar.

Ejercicio 3. Demuéstrase la siguiente identidad para los números de Bernoulli:

$$\sum_{k=0}^{2n} (-1)^k \binom{2n}{k} 2^k B_k B_{2n-k} = (1-2n) B_{2n}.$$

El desarrollo de la teoría y del aparato de las funciones generatrices condujo a la introducción no sólo de las series de potencias, sino también de los análogos del teorema de Taylor: en otras palabras, a la expresión de las funciones generatrices en términos de sus derivadas. Sea dada una función compuesta $A(t) = f(g(t))$, $g(t) = u$. Introduzcamos las designaciones para las derivadas de esta función:

$$\frac{d}{dt} = D_t, \quad D_t^n A(t) = A_n,$$

$$\frac{d}{du} = D_u, \quad D_u^n g(t) = g_n, \quad [D_u^n f(u)]_{u=g(t)} = f_n,$$

entonces tenemos en estas designaciones:

$$\begin{aligned} A_1 &= f_1 g_1, \\ A_2 &= f_1 g_2 + f_2 g_1^2, \\ A_3 &= f_1 g_3 + 3f_2 g_1 g_2 + f_3 g_1^3 \\ &\dots \dots \dots \\ A_n &= \sum_{k=1}^n f_k A_{n-k}(g_1, g_2, \dots, g_n), \end{aligned}$$

donde A_{n-k} sólo dependen de g_i ($i = 1, 2, \dots, n$) y no dependen de f_k . Eli-jamos para la función f una forma que sea próxima a las funciones genera-rices exponenciales; hallemos, además, una expresión cómoda para $A_{n,k}$. Sea

$$f(g) = \exp(ag), \quad a = \text{const.}$$

Entonces

$$\begin{aligned} f_k &= a^k \exp(ag), \quad g = g(t), \\ \sum_k A_{n-k}(g_1, g_2, \dots, g_n) a^k &= e^{-ag} D_t^n e^{ag} = A_n(a; g_1, g_2, \dots, g_n), \end{aligned}$$

donde el segundo miembro es una notación abreviada del primer miembro de la igualdad. La expresión para A_n se escribirá del modo siguiente:

$$A_n = L_f A_n(f; g_1, g_2, \dots, g_n), \quad L_f f^k = f_k, \quad k = 0, 1, \dots$$

De acuerdo con la definición de $A(t)$, agreguemos:

$$A_0 = f_0 = A(t).$$

Los polinomios $A_n(a; g_1, g_2, \dots, g_n)$ (o bien, a veces, su forma particular para $a = 1$) llevan el nombre de *polinomios de Bell*; en las designaciones de

Bell $A_n = (1; y_1, y_2, \dots, y_n) = Y(y_1, y_2, \dots, y_n) = e^{-y} D_x^n e^y$, $y = y(x)$.

Dichos polinomios resultaron ser un medio operacional importante en la matemática estadística, y luego en la teoría combinatoria.

Halleemos una relación recurrente para los polinomios de Bell. Introduzcamos la designación abreviada

$$A_n(a) = A_n(a; g_1, \dots, g_n).$$

Entonces

$$A_{n+1}(a) = e^{-ag} D^n (d e^{ag}) = e^{-ag} a D^n (e^{ag}).$$

De aquí, con arreglo a la fórmula de Leibniz sobre la diferenciación de un producto, tenemos

$$\begin{aligned} A_{n+1}(a) &= a \sum_{k=0}^n \binom{n}{k} (e^{-ag} D^{n-k} e^{ag}) D^k = a \sum_{k=0}^n \binom{n}{k} A_{n-k}(a) g^{k+1} = \\ &= L_{A(a)} L_g a g [A(a) + g]^n, \\ L_{A(a)} [A(a)]^k &= A_k(a), \quad L_g g^k = g^k, \quad k = 0, 1, 2, \dots \end{aligned} \quad (7)$$

Veamos ahora la serie

$$G(x) = \sum_{n=1}^{\infty} g_n \frac{x^n}{n!}$$

y la función generatriz exponencial para los polinomios de

$$F(x) = \sum_{n=0}^{\infty} A_n(a) \frac{x^n}{n!}$$

De (7) obtenemos

$$DF(x) = aG(x)F(x),$$

de aquí, integrando esta ecuación, obtenemos

$$F(x) = \exp(aG(x)).$$

Ahora, según el teorema polinomial, tenemos

$$A_n(a) = \sum \frac{a^k n!}{k_1! \dots k_n!} \left(\frac{g_1}{1!} \right)^{k_1} \dots \left(\frac{g_n}{n!} \right)^{k_n},$$

donde $k = k_1 + \dots + k_n$, y la suma se toma por todas las soluciones en números enteros de la ecuación $k_1 + 2k_2 + \dots + nk_n = n$. Y, por fin, obtenemos la expresión

$$A_n = A_n(f) = \sum \frac{n! f_k}{k_1! \dots k_n!} \left(\frac{g_1}{1!} \right)^{k_1} \dots \left(\frac{g_n}{n!} \right)^{k_n},$$

que se conoce como fórmula de F. di Bruno

2.4. SOBRE LAS APLICACIONES DEL MÉTODO DE FUNCIONES GENERATRICES

Según lo dicho más arriba, el método de funciones generatrices se aplica para resolver problemas de carácter enumerativo, es decir, cuando se determina el número de objetos en cierta clase. Los razonamientos combinatorios inmediatos análogos a aquellos que ya se usaron en el cap. I jugaban y siguen jugando gran papel en calidad de métodos principales, mas las funciones generatrices introdujeron consigo una mayor generalidad de juicios y ampliaron el dominio de las aplicaciones posibles del análisis combinatorio.

Históricamente las circunstancias tomaron un cariz tal que los métodos combinatorios se aplicaron, ante todo, en la teoría de las probabilidades y en la estadística matemática. Y esto es bien explicable. En la moderna teoría axiomática de las probabilidades estas mismas se interpretan sólo en relación con los espacios de los sucesos elementales. Los últimos son conceptos originarios no definibles y se interpretan como conjuntos puntuales. Si un espacio de sucesos elementales se compone de un conjunto finito o numerable de puntos, dicho espacio se denomina discreto.

Todo el aparato de investigación de los espacios discretos de sucesos elementales es, en esencia, combinatorio. Más aún, podemos hablar sobre las interpretaciones teórico-probabilísticas de una parte determinada del análisis combinatorio.

Efectivamente, las r -muestras pueden ser interpretadas con ayuda de diferentes esquemas de urna. Los casos en que se admiten repeticiones de los elementos en las muestras, corresponden a los esquemas de urna con retorno. Las funciones generatrices de la forma

$$\sum_{i=1}^s p_i t^{x_i}$$

donde $\sum_{i=1}^s p_i = 1$, $p_i \geq 0$, ($i = 1, \dots, s$) intervienen en la teoría de las probabilidades, cuando una magnitud aleatoria X (una función sobre el conjunto de sucesos elementales) puede asumir los valores x_i , $i = 1, 2, \dots, s$, con las probabilidades p_i , $i = 1, 2, \dots, s$, respectivamente. Una de las funciones generatrices más simples en la teoría de las probabilidades está asociada con el lanzamiento de una moneda y con otros sucesos que pueden tener dos resultados:

$$P(t) = q + pt, \quad p + q = 1, \quad p, q \geq 0$$

(la magnitud aleatoria toma el valor 0 con la probabilidad q , y el valor 1 con la probabilidad p). La función generatriz de lanzamiento de un dado hexaedro con las caídas equiprobables de puntos tiene la forma

$$\frac{1}{6} (t + t^2 + \dots + t^6) = \frac{t(1 - t^6)}{6(1 - t)}.$$

Sea X_1 una magnitud aleatoria que toma los valores $0, 1, 2, \dots, s$ con las probabilidades $p_0, p_1, p_2, \dots, p_s$, respectivamente; supongamos que X_2 es otra magnitud aleatoria que toma los valores $0, 1, 2, \dots, r$ con las probabilidades $p'_0, p'_1, p'_2, \dots, p'_r$, respectivamente. Sus funciones generatrices serán

$$\sum_{i=0}^s p_i t^i \quad \text{y} \quad \sum_{k=0}^r p'_k t'^k.$$

Si dos sucesos son independientes, entonces la probabilidad de su aparición simultánea es igual al producto de sus probabilidades. La correspondiente función generatriz de la magnitud aleatoria bidimensional (X_1, X_2) tiene por expresión:

$$\begin{aligned} p_0 p'_0 + (p_1 p'_0) t + (p_0 p'_1) t' + (p_1 p'_1) t t' + \dots + (p_s p'_r) t^s t'^r = \\ = \left(\sum_{i=0}^s p_i t^i \right) \left(\sum_{k=0}^r p'_k (t')^k \right). \end{aligned}$$

Si analizamos la distribución de la suma $X_1 + X_2$, entonces $t = t'$, y el producto tiene por expresión

$$\begin{aligned} \left(\sum_i p_i t^i \right) \left(\sum_k p'_k t^k \right) = p_0 p'_0 + (p_0 p'_1 + p_1 p'_0) t + \\ + (p_1 p'_1 + p_0 p'_2 + p_2 p'_0) t^2 + \dots \end{aligned}$$

Esta construcción se extiende con facilidad a las distribuciones de n magnitudes aleatorias independientes. En particular, para n pruebas independientes (con 2 resultados) la función generatriz del número de resultados de una forma determinada será

$$P(t) = (q + pt)^n, \quad p + q = 1, \quad p, q \geq 0.$$

En la teoría de las probabilidades pueden emplearse también funciones generatrices de tipo algo diferente. Por ejemplo, supongamos que cuatro partículas entran volando en una cámara, donde se encuentran centros de atracción con fuerzas proporcionales a 8:9:11:12. Diremos que la función generatriz de las probabilidades de distribución de una partícula por los centros es

$$\frac{8}{40} t_1 + \frac{9}{40} t_2 + \frac{11}{40} t_3 + \frac{12}{40} t_4.$$

En tal caso, si se asegura la independencia de las partículas, la función generatriz de las probabilidades de caída de las cuatro partículas tendrá por expresión:

$$\prod_{i=1}^4 \left(\frac{8}{40} t_1^{(i)} + \frac{9}{40} t_2^{(i)} + \frac{11}{40} t_3^{(i)} + \frac{12}{40} t_4^{(i)} \right).$$

Si nos interesa sólo cuántas partículas caen en los centros dados, la función

generadora correspondiente será

$$\left(\frac{8}{40} t_1 + \frac{9}{40} t_2 + \frac{11}{40} t_3 + \frac{12}{40} t_4 \right)^4. \quad (1)$$

Por ejemplo, la probabilidad de que las partículas caigan en los centros diferentes será igual al coeficiente de $t_1 t_2 t_3 t_4$ en (1).

Los problemas combinatorios de alojamiento en su interpretación teórico-probabilística conducen a la introducción de las nociones de momentos. Sea dada una distribución de las probabilidades, es decir, una sucesión

$$p_0, p_1, p_2, \dots; 0 \leq p_i \leq 1; i = 0, 1, 2, \dots; \sum_{i=0}^{\infty} p_i = 1.$$

Para la distribución citada pueden introducirse momentos de diferentes tipos:

a) momentos ordinarios

$$m_k = \sum_{i=0}^{\infty} i^k p_i,$$

$k = 0, 1, 2, \dots$, en particular

$$\begin{aligned} m_0 &= p_0 + p_1 + p_2 + \dots, \\ m_1 &= p_1 + 2p_2 + 3p_3 + \dots, \\ &\dots \dots \dots \\ m_k &= p_1 + 2^k p_2 + 3^k p_3 + \dots \end{aligned}$$

Los momentos ordinarios se identifican con las esperanzas matemáticas: si está dada una magnitud aleatoria X con distribución $\{f(x_i)\}$ y si la esperanza matemática de la magnitud X^r ($r > 0$) existe, ésta se llamará momento de r -ésimo orden para X :

b) momentos factoriales:

$$(m_k) = \sum_{j=0}^{\infty} (j)_k p_j,$$

donde $(j)_k = j(j-1) \dots (j-k+1)$, $k = 0, 1, 2, \dots$; la designación $(m)_k$ se debe a que $(m)_k = L_m m(m-1) \dots (m-k+1)$, donde $L_m(m') = m'$, con la particularidad de que m_i es un momento ordinario de i -ésimo orden;

c) momentos binomiales:

$$B_k = \sum_{j=0}^{\infty} \binom{j}{k} p_j, \quad k = 0, 1, \dots$$

Entre los momentos $(m)_k$ y B_k existe una relación evidente:

$$B_k = \frac{(m)_k}{k!}.$$

Es obvio que la expresión de los momentos factoriales a través de los ordinarios genera los números de Stirling de primer género, y la expresión inversa, los números de Stirling de segundo género;

d) momentos centrales:

$$M_k = L_m(m - m_1)^k, \quad L_m(m^i) = m_i, \quad i = 0, 1, 2, \dots,$$

o bien, de otro modo,

$$M_k = \sum_{j=0}^k \binom{k}{j} m_{k-j} (-m_1)^j, \quad k = 0, 1, 2, \dots$$

El más simple es el segundo momento central llamado de otro modo varianza: $M_2 = m_2 - m_1^2$. La idea principal que condujo al análisis de los momentos centrales consiste en que en lugar de la misma magnitud aleatoria se examinan sus desviaciones del valor medio.

He aquí algunos tipos de las funciones generatrices que se usan con mayor frecuencia en la teoría de las probabilidades:

a) para la distribución de las probabilidades:

$$P(t) = \sum_k p_k t^k;$$

b) para los momentos ordinarios;

$$m(t) = L_m \exp(mt) = \sum_k m_k \frac{t^k}{k!}; \quad L_m(m^k) = m_k, \quad k = 0, 1, \dots;$$

c) para los momentos factoriales:

$$(m)(t) = \sum_k (m)_k \frac{t^k}{k!};$$

d) para los momentos binomiales

$$B(t) = \sum_k B_k t^k$$

e) para los momentos centrales:

$$M(t) = L_M \exp(Mt) = \sum_k M_k \frac{t^k}{k!}, \quad L_M(M^k) = M_k, \quad k = 0, 1, \dots$$

Las operaciones de carácter formal sobre diferentes funciones generatrices conducen al descubrimiento de varias relaciones útiles, por ejemplo:

1) $m(t) = P(e^t)$. En efecto,

$$m(t) = \sum_k m_k \frac{t^k}{k!} = \sum_k \frac{t^k}{k!} \sum_j j^k p_j = \sum_j p_j \sum_k \frac{(tj)^k}{k!} = \sum_j p_j e^{tj} = P(e^t);$$

2) $(m)(t) = P(1+t)$, lo que se ve de las siguientes igualdades:

$$\begin{aligned}(m)(t) &= \sum_k (m)_k \frac{t^k}{k!} = \sum_k \sum_j (j)_k p_j \frac{t^k}{k!} = \\ &= \sum_j p_j \sum_k \binom{j}{k} t^k = \sum_j p_j (1+t)^j = P(1+t).\end{aligned}$$

2.5. TEORÍA DE REDFIELD—POLYA

El tipo de funciones generatrices depende de las condiciones concretas en que se plantean los problemas y su construcción es, en cierto grado, un arte. Durante mucho tiempo no existía un enfoque regular respecto a la construcción de las funciones generatrices. En este sentido han hecho pasos considerables Redfield y Polya, que elaboraron el método de construcción de las funciones generatrices para objetos combinatorios no equivalentes de tipo bastante general. A los objetos enumerados se les atribuían pesos, mientras que la noción de equivalencia se introducía por medio de un grupo de sustituciones. Aclaremos esta cuestión detalladamente.

Veamos un grupo G de sustituciones de los elementos de un n -conjunto, es decir, un subgrupo de un grupo simétrico S_n . Cada sustitución engendra la partición del n -conjunto en ciclos, es decir, en subconjuntos cuyos elementos son cíclicamente permutables: la cantidad de elementos en un subconjunto se denomina longitud del ciclo. Supongamos, por ejemplo, que la sustitución g tiene k_1 ciclos de longitud 1, k_2 ciclos de longitud 2, etc. Entonces suele decirse que la sustitución posee estructura cíclica $(k_1, k_2, \dots)$ y se le pone en correspondencia un monomio $p(g) = t_1^{k_1} t_2^{k_2} \dots$. Se llama *índice cíclico* del grupo G la media de tales monomios tomada de todas las $g \in G$:

$$P_G(t_1, t_2, \dots) = |G|^{-1} \sum_{g \in G} t_1^{k_1} t_2^{k_2} \dots$$

EJEMPLO. Basándonos en la fórmula (4) del § 1.3 para el número de sustituciones que poseen estructura cíclica dada, encontramos el índice cíclico del grupo simétrico S_n para todas las sustituciones del n -conjunto:

$$\begin{aligned}P_{S_n}(t_1, t_2, \dots, t_n) &= \frac{1}{n!} \sum \frac{n!}{1^{k_1} k_1! \dots n^{k_n} k_n!} t_1^{k_1} \dots t_n^{k_n} = \\ &= \sum \frac{1}{k_1! \dots k_n!} \left(\frac{t_1}{1}\right)^{k_1} \dots \left(\frac{t_n}{n}\right)^{k_n},\end{aligned}$$

donde la sumación se realiza respecto de todas las soluciones en números enteros no negativos de la ecuación $1 \cdot k_1 + 2 \cdot k_2 + \dots + n k_n = n$, o bien, hablando de otro modo, respecto de todas las particiones del número n .

A todo grupo de sustituciones se le puede asignar su índice cíclico de

un modo único. Una afirmación contraria no será, sin embargo, cierta. En 1937 Polya construyó dos grupos no isomorfos de sustituciones de orden p^3 ($p > 2$, primo) con iguales índices cíclicos.

Sean dados un conjunto finito D y un grupo finito G con su aplicación homomorfa π en un grupo simétrico de sustituciones del conjunto D , es decir, a todo $g \in G$ le corresponde una sustitución π_g del conjunto D , y $\pi_{gg'} = \pi_g \pi_{g'}$ (para cualesquiera $g, g' \in G$). No se requiere que la aplicación π sea un encaje.

La equivalencia de los elementos del conjunto D se introduce por medio del grupo G :

$$d_1 \sim d_2 \quad (d_1, d_2 \in D),$$

si existe un elemento $g \in G$ tal que $\pi_g d_1 = d_2$. Todos los requerimientos de equivalencia se cumplen en esta definición. En efecto:

a) $d \sim d$ para todos los $d \in D$, puesto que π_e , donde e es un elemento unidad de G , es una sustitución idéntica;

b) de $d_1 \sim d_2$ se deduce $d_2 \sim d_1$, por cuanto de la condición de homomorfismo se desprende que si $g' = g^{-1}$, entonces $\pi_{g'} = (\pi_g)^{-1}$;

c) de $d_1 \sim d_2$ y $d_2 \sim d_3$ se deduce que $d_1 \sim d_3$, puesto que si $\pi_g d_1 = d_2$ y $\pi_{g'} d_2 = d_3$, entonces $\pi_{g'g} d_1 = \pi_{g'}(\pi_g d_1) = \pi_{g'} d_2 = d_3$.

En virtud de la definición de equivalencia, el conjunto D resulta ser partido, con ayuda del grupo G , en clases de equivalencia (conjuntos transitivos). Los elementos son equivalentes, si y sólo si integran una misma clase de equivalencia.

Lema de Bernsáid. El número de clases de equivalencia definidas por el grupo finito G que actúa sobre el conjunto finito D es igual a

$$\frac{1}{|G|} \sum_{g \in G} \psi(g),$$

donde $|G|$ es el número de elementos en G ; $\psi(g)$, el número de elementos en D , que son invariantes respecto de π_g , es decir, tales que $\pi_g d = d$, mientras que la sumación se hace de todos los elementos $g \in G$.

Esta afirmación se anunció por primera vez en las obras de Cauchy, Frobenius y Bernsáid.

Demostración. Veamos todos los pares (g, d) , para los cuales $\pi_g d = d$ ($g \in G, d \in D$) entonces:

a) podemos fijar g y contar cuántos d existen, para los cuales $\pi_g d = d$;

b) para cada d fijo calculemos todos los g , para los cuales $\pi_g d = d$; designemos los números obtenidos con $\eta(d)$. Está claro que

$$\sum_{d \in D} \eta(d) = \sum_{g \in G} \psi(g).$$

Los elementos $g \in G$ que fijan d forman un subgrupo G_d del grupo G :

$$|G_d| = \eta(d).$$

Si tomamos otro elemento de la misma clase de equivalencia $d_1 \sim d$, entonces el número de elementos g , para los cuales $\pi_g d = d_1$, será también igual a $|G_d|$. Efectivamente, se tiene un elemento $h \in G$, para el cual $\pi_h d_1 = d$, y, por tanto, de que $\pi_g d = d_1$ se deduce que $hg \in G_d$. Así pues, el grupo G puede ser partido en subconjuntos disjuntos en cada uno de los cuales habrá $|G_d|$ elementos (un subgrupo G_d en cada clase contigua izquierda del grupo G): cada subconjunto corresponde exactamente a un solo elemento de la clase de equivalencia, en la que figura el elemento d . De aquí, para hallar $\eta(d)$, se debe dividir $|G|$ por el número de elementos de la clase de equivalencia a la que pertenece d . Sumamos: a) $\sum \eta(d) = |G|$, si la suma se toma de todos los d que integran una misma clase de equivalencia; b) la suma $\sum_{d \in D} \eta(d)$ es igual al producto de $|G|$ por el número de clases de equivalencia, de donde se desprende el resultado que tratábamos de demostrar.

Hemos introducido, pues, el concepto de equivalencia para los elementos a través de un grupo de sustituciones y hemos deducido la fórmula para el número de clases de equivalencia.

Introduzcamos la relación de equivalencia para las aplicaciones. Sean dados los conjuntos finitos D y R . Designemos el conjunto de todas las aplicaciones de D en R con R^D . El número de estas aplicaciones es igual a $|R|^{|D|}$, puesto que para cada $d \in D$ existe $|R|$ posibilidades independientes para la imagen. Supongamos que se dispone de un grupo G de sustituciones del conjunto D . Las aplicaciones f_1 y f_2 son equivalentes ($f_1 \sim f_2$), si existe una sustitución $g \in G$ tal que $f_1(gd) = f_2(d)$ para todos los $d \in D$, o bien $f_1 g = f_2$. Las condiciones de equivalencia son: a) $f \sim f$; b) si $f_1 \sim f_2$, entonces $f_2 \sim f_1$; c) si $f_1 \sim f_2$ y $f_2 \sim f_3$, entonces $f_1 \sim f_3$ se cumplen todas. La primera condición se deduce de la existencia en G de una sustitución idéntica; la segunda, de que junto con g figura en G también g^{-1} ; la tercera, de que si $g_1 \in G$ y $g_2 \in G$, entonces $g_1 g_2 \in G$.

La equivalencia de las aplicaciones introducida de este modo parte el conjunto R^D en clases de equivalencia (*modelos*).

Introduzcamos ahora los pesos. Estos darán la posibilidad de enumerar las clases especiales de los objetos combinatorios. Al principio, a todo elemento $r \in R$ se le asigna un peso $\omega(r)$, donde $\omega(r)$ es un elemento de anillo conmutativo sobre el campo de números racionales. El peso de la aplicación $W(f)$, donde $f \in R^D$, se define como un producto

$$W(f) = \prod_{d \in D} \omega(f(d)),$$

donde $f(d)$ es la imagen del elemento $d \in D$ al realizarse la aplicación f , y $\omega(f(d))$, su peso. Las aplicaciones equivalentes tienen pesos iguales. En efecto, si $f_1 g = f_2$ ($g \in G$), entonces

$$\prod_{d \in D} \omega(f_1(d)) = \prod_{d \in D} \omega(f_1(gd)) = \prod_{d \in D} \omega(f_2(d)).$$

La igualdad es obvia, puesto que tanto el primer producto, como el segundo, difieren sólo en el orden de los factores. Por cuanto todas las aplicaciones f de la clase de equivalencia F poseen un mismo peso, asignemos este peso a toda la clase F , designándolo con $W(F)$. La suma de pesos

$$\sum_{r \in R} \omega(r)$$

de los elementos del conjunto R se llamará (siguiendo a de Bruijn) *inventario* (inventory) del conjunto R y se denotará con $\text{inv } D$.

Ahora podemos calcular el inventario del conjunto R^D de aplicaciones

$$\text{inv } R^D = \sum_f W(f) = \left(\sum_{r \in R} \omega(r) \right)^{|D|} = (\text{inv } R)^{|D|}.$$

Mostremos que esta afirmación es cierta. En el segundo miembro tenemos $|R|^{|D|}$ términos. Se puede establecer una correspondencia entre $|D|$ factores y los elementos del conjunto D . La elección de un término de cada factor y la formación de un término del desarrollo se interpretan como la aplicación $f: D \rightarrow R$. Pero, para cada f tenemos $W(f) = \prod_{d \in D} \omega(f(d))$. Todo el producto es igual a la suma de todos los $W(f)$, lo que se requería demostrar.

Si consideramos el conjunto S de aplicaciones, constantes sobre los subconjuntos $D_1, \dots, D_k$, que forman la partición del conjunto D , entonces

$$\text{inv } S = \prod_{i=1}^k \sum_{r \in R} (\omega(r))^{|D_i|}, \quad (1)$$

lo que se demuestra mediante un razonamiento análogo al aducido más arriba.

Ahora, introducidos todos los conceptos indispensables, podemos enunciar el teorema básico. Se trata en él sobre el paso de los pesos de las funciones a los de los modelos (de las clases de equivalencia).

Teorema de Polya. Sean D y R los conjuntos finitos dados, y G , un grupo de sustituciones de los elementos de D : supongamos, además, que a los elementos $r \in R$ están asignados los pesos $\omega(r)$; a las aplicaciones $f \in R^D$ y a los modelos (clases) F , se les asignan los pesos $W(f)$ y $W(F)$, respectivamente. Entonces

$$\sum_F W(F) = P_G \left(\sum_{r \in R} \omega(r), \sum_{r \in R} (\omega(r))^2, \sum_{r \in R} (\omega(r))^3, \dots \right),$$

donde P_G es el índice cíclico del grupo G . En un caso particular, cuando todos los pesos son iguales a uno, el número de clases de equivalencia será

$$P_G(|R|, |R|, \dots).$$

Demostración. Elijamos de R^D , es decir, del conjunto de todas las aplicaciones $f: D \rightarrow R$, un conjunto S de aquellas aplicaciones, cuyos pesos son iguales, digamos, a ω . Para $f \in S$ tenemos: $W(f) = \omega$, y el número de clases

de equivalencia para el conjunto S es igual a

$$\frac{1}{|G|} \sum_{g \in G} \psi_{\omega}(g),$$

donde $\psi_{\omega}(g)$ denota el número de aplicaciones f , para las cuales

$$W(f) = \omega, \quad f = fg \text{ (o bien } fg^{-1} = f).$$

En efecto, se ha demostrado que si $g \in G$ y $f_1 = f_2 g$, entonces f_1 y f_2 son de peso igual. Quiere decir, que si $f_1 \in S$, también $f_1 g^{-1} \in S$. De este modo, a cada $g \in G$ corresponde una aplicación π_g del conjunto S sobre sí mismo que se determina por la correlación

$$\pi_g f = fg^{-1}.$$

Si dos elementos f_1 y f_2 del conjunto S son equivalentes, esta afirmación es equivalente a la siguiente: la existencia del elemento $g \in G$, para el cual $\pi_g f_2 = f_1$, es equivalente a la existencia del elemento $g \in G$, para el cual $f_2 = f_1 g$. Ahora resta por alegar el lema de Bernsaid para convencerse de que el número de clases de equivalencia está determinado correctamente.

Todas las clases de equivalencia que integran S tienen un peso igual a ω . Por consiguiente, si multiplicamos

$$\frac{1}{|G|} \sum_{g \in G} \psi_{\omega}(g)$$

por ω y sumamos respecto de todos los ω , obtendremos

$$\sum_F W(F) = \frac{1}{|G|} \sum_{\omega} \sum_{g \in G} \psi_{\omega}(g) \omega.$$

Pero, es evidente que

$$\sum_{\omega} \psi_{\omega}(g) \omega = \sum_f^{(g)} W(f),$$

donde la sumación se realiza por todas las $f \in R^D$ que satisfacen la igualdad $f = fg$; quiere decir

$$\sum_F W(F) = \frac{1}{|G|} \sum_{g \in G} \sum_f^{(g)} W(f). \quad (2)$$

Calculemos ahora $\sum_f^{(g)} W(f)$. La sustitución g divide D en ciclos. La condición $f = fg$ significa que

$$f(d) = f(gd) = f(g^2 d) = \dots,$$

es decir, $f = \text{const}$ para cada ciclo. Es cierta también la afirmación contra-

ria: toda aplicación f , constante en cada ciclo, satisface la igualdad $fg = f$, puesto que $g(d)$ y d pertenecen siempre a un mismo ciclo. De modo que si $D_1, D_2, \dots, D_k$ son los ciclos de g , entonces, según (1), tenemos

$$\sum_f^{(g)} W(f) = \prod_{i=1}^k \sum_{r \in R} (\omega(r))^{1_{D_i}}.$$

Supongamos que la sustitución g es de estructura cíclica $(b_1, b_2, \dots)$, entonces

$$\sum_f^{(g)} W(f) = \left(\sum_{r \in R} \omega(r) \right)^{b_1} \left(\sum_{r \in R} (\omega(r))^2 \right)^{b_2} \dots$$

Al sustituir esta expresión en (2), obtenemos

$$\sum_f W(f) = \frac{1}{|G|} \sum_{g \in G} \left(\sum_{r \in R} \omega(r) \right)^{b_1} \left(\sum_{r \in R} (\omega(r))^2 \right)^{b_2} \dots$$

lo que se trataba de demostrar.

Este teorema fue enunciado por Polya en 1937. Más tarde se notó que el mismo teorema, pero en una forma algo diferente, fue publicado en 1927 por J. Redfield (véase en [12]). Sin embargo, debido a una costumbre ya arraigada el teorema conserva el nombre de Polya.

Posteriormente dicho teorema fue demostrado basándose en suposiciones generales. Según lo observado anteriormente, para las aplicaciones $f: D \rightarrow R$ la equivalencia se introducía por intermedio del grupo G de sustituciones de los elementos $d \in D$. Si añadimos el segundo grupo H de sustituciones de los elementos $r \in R$, la introducción de la equivalencia con ayuda de ambos grupos G y H se realiza del modo siguiente: dos funciones (dos aplicaciones) $f_1 \in R^D$ y $f_2 \in R^D$ son equivalentes: $f_1 \sim f_2$, si existen elementos $g \in G$ y $h \in H$ tales que $f_1 g = h f_2$, es decir,

$$f_1(gd) = h f_2(d) \text{ para todos } d \in D.$$

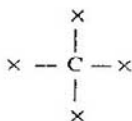
Entre otras generalizaciones del teorema de Polya indiquemos aquella que está ligada con el enlace de los grupos. Sean los grupos de sustituciones G y H para los conjuntos S y T , respectivamente. En el producto directo de los conjuntos $S \times T$ formemos las sustituciones de un tipo especial. A saber, elegimos $g \in G$; para cada $s \in S$ elegimos un elemento $h_s \in H$. Estos elementos definen la sustitución:

$$(s, t) \rightarrow (gs, h_s t), \quad s \in S, \quad t \in T.$$

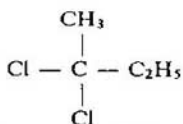
En total hay $|G| \cdot |H|^{|\mathcal{S}|}$ de tales sustituciones; ellas forman precisamente un grupo denominado por Polya como enlace de los grupos (Gruppenkranz).

La explicación acertada de los teoremas de Polya provista de ejemplos se da en [10]. El desarrollo ulterior ha enriquecido poco la teoría de Redfield — Polya y no la cambió en principio.

Como conclusión, aduzcamos un ejemplo no complejo que explica la aplicación del teorema de Polya. Se trata de la enumeración de los isómeros de las moléculas orgánicas de una estructura dada. Se consideran las moléculas del tipo:



donde C es un átomo de carbono, y en los lugares marcados con cruces pueden encontrarse: CH₃ (metilo), C₂H₅ (etilo), H (hidrógeno) y Cl (cloro). Por ejemplo,



(diclorobutano).

De modelo matemático de estas moléculas sirve un tetraedro en cuyo centro se dispone el átomo de carbono. El problema de enumeración de las moléculas se interpretará como problema sobre el número de clases de equivalencia *D* (para cuatro vértices):

$$f : D \rightarrow R = \{\text{CH}_3, \text{C}_2\text{H}_5, \text{H}, \text{Cl}\}.$$

El grupo *G* será un grupo de rotaciones del tetraedro compuesto por: una permutación idéntica de los vértices; ocho rotaciones en 120° alrededor de los vértices; tres rotaciones en 180° alrededor de los ejes que pasan por los centros de las aristas que no se intersectan. Entonces

$$P_G(x_1, x_2, x_3) = \frac{1}{12} (x_1^4 + 8x_1x_3 + 3x_2^2).$$

Supongamos que todos los pesos son iguales a la unidad y obtendremos el número total de moléculas:

$$P_G(4, 4, 4) = 36.$$

Sustituyamos los tipos particulares de las moléculas de la estructura dada: supongamos, por ejemplo, que se debe calcular el número de tales moléculas, en las que no interviene el átomo de hidrógeno. En este caso para CH₃, C₂H₅ y Cl los pesos serán iguales a la unidad, y para H, a cero:

$$P_G(3, 3, 3) = \frac{1}{12} (3^4 + 8 \cdot 3 \cdot 3 + 3 \cdot 3^2) = 15 \text{ moléculas buscadas.}$$

Para poder clasificar las demás 21 moléculas atribuyamos, como lo hicimos antes, los pesos unidad a CH₃, C₂H₅ y Cl, y el peso H al átomo de hidrógeno:

$$P_G(H + 3, H^2 + 3, H^3 + 3) = \frac{1}{12} ((H + 3)^4 + 8(H + 3)(H^3 + 3) + 3(H^2 + 3)^2) = H^4 + 3H^3 + 6H^2 + 11H + 15.$$

Por consiguiente, existen: 1 molécula de CH_4 (metano); 3 moléculas con 3 átomos de H; 6 moléculas con 2 átomos de H; 11 moléculas con 1 átomo de H; 15 moléculas que no tienen átomos de H.

La teoría de las funciones generatrices es la parte del análisis combinatorio que mayor avance tiene en los aspectos teórico y práctico para la resolución de los problemas prácticos. La composición y las posibilidades de esta teoría ya están bien determinadas.

Capítulo 3 Métodos lógicos

En este capítulo describiremos procedimientos lógicos característicos que constituyen la base de muchas demostraciones combinatorias.

3.1. MÉTODO DE INCLUSIONES Y EXCLUSIONES¹⁾

Sean dados un n -conjunto S de ciertos elementos y un N -conjunto de propiedades $p_1, p_2, \dots, p_N$, con la particularidad de que los elementos del conjunto pueden poseer dichas propiedades y pueden no tenerlas. Se requiere hallar el número de elementos que no poseen ninguna de las propiedades mencionadas.

Elijamos una r -muestra de propiedades $(p_{i_1}, p_{i_2}, \dots, p_{i_r})$. Designemos con $n(p_{i_1}, p_{i_2}, \dots, p_{i_r})$ el número de elementos del conjunto S , cada uno de los cuales posee todas las propiedades elegidas. Denotemos con $\bar{p}_i$ la ausencia en los elementos de la propiedad p_i . Así pues, el número de elementos que, digamos, poseen las propiedades p_1, p_3, p_5 , y no poseen las propiedades p_2, p_4, p_6 se escribirá en la forma $n(p_1, \bar{p}_2, p_3, \bar{p}_4, p_5, \bar{p}_6)$.

Examinemos al principio dos casos sencillos:

a) se tiene sólo una propiedad p ; entonces, evidentemente,

$$n(\bar{p}) = n - n(p);$$

b) se tiene un número finito de propiedades $p_1, p_2, \dots, p_N$ que no son compatibles; tenemos, evidentemente

$$n(\bar{p}_1, \bar{p}_2, \dots, \bar{p}_N) = n - \sum_{i=1}^N n(p_i).$$

Pasemos a un planteamiento más general del problema, cuando los elementos del conjunto pueden poseer combinaciones de las propiedades compatibles. En este caso tiene lugar el

Teorema 1. Si están dados un n -conjunto de elementos y un N -conjunto de propiedades p_i ; $i = 1, 2, \dots, N$, entonces

$$\begin{aligned} n(\bar{p}_1, \bar{p}_2, \dots, \bar{p}_N) = & n - \sum_{i=1}^N n(p_i) + \sum_{1 \leq i < j \leq N} n(p_i, p_j) - \\ & - \sum_{1 \leq i < j < k \leq N} n(p_i, p_j, p_k) + \dots + (-1)^N n(p_1, p_2, \dots, p_N). \end{aligned} \quad (1)$$

¹⁾ Se conoce también bajo la denominación de método de criba, método lógico, método simbólico y principio de clasificación cruzada.

Demostración. Con el fin de obtener elementos que no posean ninguna de las propiedades mencionadas, hace falta excluir del n -conjunto los elementos que tienen la propiedad p_1 , luego, los elementos poseedores de la propiedad p_2 , etc., es decir, $\sum_i n(p_i)$ elementos. En este caso, sin embargo,

los elementos que poseen dos propiedades, digamos p_1 y p_2 resultaron excluidos dos veces (al principio, como poseedores de la propiedad p_1 , y a continuación, como poseedores de la propiedad p_2). Quiere decir, que es necesario devolver todos los conjuntos, cuyos elementos poseen dos propiedades, o sea hay que añadir $\sum_{1 \leq i < j \leq N} n(p_i, p_j)$ elementos. Mas, en tal caso los

elementos que poseen tres propiedades, digamos p_1 , p_2 y p_3 quedan incluidos, por lo cual se debe sustraer $\sum_{1 \leq i < j < k \leq N} n(p_i, p_j, p_k)$, elementos.

Siguiendo razonando de un modo análogo, obtendremos un algoritmo para calcular $n(\bar{p}_1, \bar{p}_2, \dots, \bar{p}_N)$ el cual consiste en rechazar y devolver alternativamente los subconjuntos. A esto se debe precisamente una de las denominaciones del método: el de inclusiones y exclusiones. Vamos a seguir este método en adelante.

A parte de los razonamientos sencillos aducidos más arriba, podemos realizar la demostración por inducción respecto de N . El teorema es válido para $N = 1$;

$$n(p) = n - n(p).$$

De acuerdo con la formulación del teorema, podemos escribir también:

$$n(p_1, p_2, \dots, p_N) = n(\bar{p}_1, p_2, \dots, \bar{p}_{N-1}) - n(\bar{p}_1, \bar{p}_2, \dots, \bar{p}_{N-1}, p_N).$$

Supongamos que el teorema es válido para $N - 1$ propiedades, es decir,

$$n(p_1, \bar{p}_2, \dots, \bar{p}_{N-1}) = n - \sum_i n(p_i) + \sum_{i < j} n(p_i, p_j) - \dots \\ \dots + (-1)^{N-1} n(p_1, p_2, \dots, p_{N-1}).$$

Pasemos al caso en que se tienen N propiedades. Apliquemos la correlación obtenida para el número $n(\bar{p}_1, \bar{p}_2, \dots, \bar{p}_{N-1}, \bar{p}_N)$:

$$n(\bar{p}_1, \bar{p}_2, \dots, \bar{p}_{N-1}, p_N) = n(p_N) - \sum_i n(p_i, p_N) + \dots \\ \dots + (-1)^{N-1} n(p_1, p_2, \dots, p_N).$$

Al sustraer esta igualdad de la anterior, llegamos a la afirmación del teorema.

El carácter de la demostración es tal que puede ser utilizada para cualquier combinación de las propiedades. En el primer miembro de la igualdad demostrada puede figurar no sólo $n(\bar{p}_1, \bar{p}_2, \dots, \bar{p}_N)$, sino también, por ejemplo, $n(p_1, \bar{p}_2, p_3, \bar{p}_4)$. En este caso el teorema se enuncia respecto a una totalidad de las propiedades p_2 y p_4 , con el cumplimiento obligatorio de las

propiedades p_1 y p_3 , del modo siguiente:

$$n(p_1, p_3, \bar{p}_2, \bar{p}_4) = n(p_1, p_3) - n(p_1, p_3, p_2) - \\ - n(p_1, p_3, p_4) + n(p_1, p_3, p_2, p_4).$$

El método se hace más complicado cuando se introducen los pesos de los elementos. Al igual que en el cap. 2, no habrá limitaciones o precisiones algunas para el concepto de peso. Para nosotros los pesos son características numéricas de los elementos de los conjuntos definidas por las condiciones del problema.

Así pues, sea dado un n -conjunto S y supongamos que a cada elemento $s_i \in S$, $i = 1, 2, \dots, n$, se le atribuye el peso $V(s_i)$. Del N -conjunto de propiedades $p_1, p_2, \dots, p_N$ elijamos una r -muestra $p_{i_1}, \dots, p_{i_r}$ y designemos con $V(p_{i_1}, p_{i_2}, \dots, p_{i_r})$ la suma de pesos de los elementos que poseen todas las r propiedades elegidas. La suma de los pesos extendida a todas las r -muestras posibles de propiedades la denotaremos con

$$\sum V(p_{i_1}, p_{i_2}, \dots, p_{i_r}) = V(r).$$

Para el caso de $r = 0$ el símbolo correspondiente $V(0)$ designará la suma de pesos de todos los elementos del conjunto S .

El teorema antecedente se enuncia en este caso del modo siguiente.

Teorema 2. Si están dados el n -conjunto S , cada elemento del cual tiene un peso, y el N -conjunto de propiedades, entonces la suma $V_N(0)$ de pesos de los elementos que no poseen ninguna de las propiedades dadas se determina por la fórmula

$$V_N(0) = V(0) - V(1) + V(2) - \dots + (-1)^N V(N).$$

Ha de notarse que el teorema 2 generaliza el teorema 1. Si todos los elementos $s_i \in S$ son de peso unitario, entonces la suma de los pesos será igual a la suma de los sumandos en la suma. En este caso $V(0) = N$ y $V_N(0)$ es igual al número de elementos del conjunto S que no tienen ninguna de las N propiedades; la fórmula que se obtiene en tal caso es precisamente la fórmula (1).

El teorema 2 puede ser, a su vez, generalizado.

Teorema 3. La suma de pesos de los elementos que poseen exactamente r propiedades de las $p_1, p_2, \dots, p_N$ se determina por la fórmula

$$V_N(r) = V(r) - \binom{r+1}{1} V(r+1) + \binom{r+2}{2} V(r+2) - \dots \\ \dots + (-1)^{N-r} \binom{N}{N-r} V(N),$$

o bien, lo que es lo mismo,

$$V_N(r) = V(r) - \binom{r+1}{r} V(r+1) + \binom{r+2}{r} V(r+2) - \dots$$

$$\dots + (-1)^{N-r} \binom{N}{r} V(N).$$

Los teoremas 2 y 3 se demuestran igual que el teorema 1. Se propone que el lector mismo realice estas demostraciones o se dirija al § 8.3, donde se aducen generalizaciones ulteriores del método de inclusiones y exclusiones.

Estudiemos, a título de ejemplo, un problema sobre desórdenes, llamado también problema de encuentros. Sea un conjunto ordenado finito de números $1, 2, 3, \dots, n$. Para estos números pueden formarse las permutaciones $a_1, a_2, \dots, a_n$. El número de todas las permutaciones es $s = n!$. Entre las permutaciones citadas hay tales, donde ninguno de los elementos conserva su lugar original: $a_i \neq i, i = 1, 2, \dots, n$. Las permutaciones de esta índole se llaman desórdenes. ¿Cuántos desórdenes existen?

Un conjunto de n elementos se examina con relación al conjunto de propiedades de los elementos de quedarse en su lugar: $p_i \sim \{a_i = i \mid i = 1, 2, \dots, n\}$. Es evidente que si s elementos quedan fijados en sus lugares, el número $N(s)$ de las respectivas permutaciones es igual a $(n-s)!$. El número de desórdenes en tal caso se determina con ayuda del método de inclusiones y exclusiones:

$$N(0) = n! - \binom{n}{1}(n-1)! + \binom{n}{2}(n-2)! - \dots + (-1)^s \times \\ \times \binom{n}{s}(n-s)! + \dots = n! \left(1 - 1 + \frac{1}{2!} - \frac{1}{3!} + \dots + \frac{(-1)^n}{n!} \right),$$

lo que representa un número entero, el más próximo a $n!e^{-1}$.

Si se trata no de los desórdenes, sino de un número de permutaciones, en las que quedan en sus lugares s elementos, entonces

$$N(s) = \frac{n!}{s!} \left(1 - 1 + \frac{1}{2!} - \frac{1}{3!} + \dots + (-1)^{n-s} \frac{1}{(n-s)!} \right).$$

Es fácil observar el desarrollo de los razonamientos durante la resolución: entre n elementos se eligen s elementos inmóviles, sirviéndose para ello de $\binom{n}{s}$ métodos; a continuación, se realiza la multiplicación (según la regla del producto) por el número de desórdenes que se encuentran entre los $(n-s)$ elementos restantes.

Demos a conocer la interpretación teórico-probabilística del método de inclusiones y exclusiones. El conjunto de elementos se interpreta en estas circunstancias como un espacio discreto de sucesos elementales, simultáneos o incompatibles (excluyentes), $a_1, a_2, \dots, a_n$, y los pesos de los elementos, como probabilidades. La probabilidad de la aparición simultánea de los sucesos $a_{i_1}, a_{i_2}, \dots, a_{i_k}$ se designará con $P(a_{i_1}, \dots, a_{i_k})$. Para el caso en que los sucesos no aparecen se emplea la llamada ecuación de Poincaré:

$$P(\bar{a}_1, \bar{a}_2, \dots, \bar{a}_n) = [1 - P(a_1)][1 - P(a_2)] \dots [1 - P(a_n)],$$

en la cual podemos abrir paréntesis, rigiéndonos por la regla: $P(a_i)P(a_j) = P(a_i, a_j)$. Esta ecuación es una generalización de la fórmula para las probabilidades de los sucesos independientes aplicada al caso de los sucesos dependientes y simultáneos. La fórmula general para el método de inclusiones y exclusiones tiene por expresión:

$$n(p_1, p_2, \dots, p_n) = n - \sum_i n(p_i) + \sum_{i,j} n(p_i, p_j) - \dots \\ \dots + (-1)^n n(p_1, p_2, \dots, p_n),$$

con la particularidad de que la fórmula se modifica: ambos miembros se dividen por n , para obtener magnitudes no absolutas, sino relativas, es decir, probabilidades

$$n^{-1}n(p_1, p_2, \dots, p_n) = 1 - n^{-1}\sum_i n(p_i) + n^{-1}\sum_{i,j} n(p_i, p_j) - \dots \\ \dots + (-1)^n n^{-1}n(p_1, p_2, \dots, p_n).$$

Más abajo señalaremos como se aplica el método de inclusiones y exclusiones al cálculo de los permanentes (§ 4.4) y al cálculo de los valores de las funciones especiales que figuran en la teoría de los números (§ 8.3).

3.2. SISTEMAS DE REPRESENTANTES DE LOS CONJUNTOS

Se examina aquí uno de los accesos combinatorios a la característica de la estructura de los conjuntos finitos. Ya por denominación se puede concebir que la idea principal consiste en sustituir el sistema de conjuntos por la reunión de sus representantes.

El planteamiento de los problemas de este tipo y los métodos de su resolución dependen de los requerimientos que dichos representantes deben satisfacer.

Sistemas de representantes distintos (s.r.d.). Sean un n -conjunto S y el conjunto $P(S)$ de todos los subconjuntos suyos. Supongamos que $M = (S_1, S_2, \dots, S_m)$ es una m -muestra de $P(S)$, y $a = (a_1, a_2, \dots, a_m)$, cierta m -muestra de S . Si a la muestra M se le puede poner en correspondencia (no forzosamente de una manera unívoca) una muestra a tal que los elementos $a_i, i = 1, 2, \dots, m$, sean distintos dos a dos y, además, $a_i \in S_i, i = 1, 2, \dots, m$, suele decirse que el elemento a_i representa el conjunto S_i , y toda la muestra $(a_1, a_2, \dots, a_m)$ se denomina sistema de representantes distintos (en forma abreviada, s.r.d.) para M . Observemos a la vez que si $i \neq j$, entonces $a_i \neq a_j$, incluso si $S_i = S_j$. Si un conjunto aparece varias veces, debe tener cada vez un representante que sea distinto de los demás.

Resulta seguidamente que el s.r.d. puede existir no para todas las colecciones de conjuntos. Si en un sistema finito los conjuntos no son vacíos y no se intersecan, el s.r.d., obviamente, existe. Tomemos un caso más complejo. Por ejemplo, si $S = (a, b, c, d, e)$, y M es una colección de cuatro conjun-

tos $S_1 = (a, b, c, d)$, $S_2 = (a, b, e)$, $S_3 = S_4 = (b, e)$, entonces existen dos s.r.d.: (c, a, b, e) y (b, a, e, d) . Pero, apenas cambiamos uno de los subconjuntos, por ejemplo, tomamos $S'_2 = (b, e)$ en lugar de S_2 y ya no podremos obtener ningún s.r.d. La respuesta a la pregunta de si existe o no un s.r.d. para una familia dada de conjuntos la da el teorema de P. Hall enunciado por éste no más tarde de 1935. El teorema formula las condiciones necesarias y suficientes para la existencia del s.r.d.

Teorema de P. Hall. Los subconjuntos $S_1, S_2, \dots, S_m$ tienen un s.r.d., cuando y sólo cuando la reunión de cualesquiera k de estos conjuntos contiene no menos de k elementos. Dicho de otro modo, el s.r.d. para $S_1, S_2, \dots, S_m$ existe, si y sólo si $S_{i_1} \cup S_{i_2} \cup \dots \cup S_{i_k}$ consta por lo menos de k elementos, siendo en este caso $k = 1, 2, \dots, m$, y $(i_1, i_2, \dots, i_m)$, cualquier k -muestra de $1, 2, \dots, m$.

Demostración. La necesidad es casi evidente, puesto que la existencia del s.r.d. asegura la presencia de un número necesario de elementos en calidad de representantes distintos. En lo que se refiere a la suficiencia, aduzcamos una formulación perfeccionada que nos da, además, la cota inferior para el número de los propios s.r.d.

Teorema. Supongamos que una familia $M = (S_1, S_2, \dots, S_m)$ satisface la condición necesaria para la existencia de un s.r.d. y que cada uno de los conjuntos $S_1, S_2, \dots, S_m$ consta por lo menos de t elementos. Entonces; a) si $t \leq m$, M tiene no menos de $t!$ s.r.d.; b) si $t > m$, la familia M tiene no menos de $t!/(t-m)!$ s.r.d.

La *demostración* la realizaremos por inducción respecto de m . Para $m = 1$ (e, incluso, para $m = 2$) el teorema es evidente. Demostremos que es lícito para cualquier m finito, partiendo de su validez para $m' < m$.

Veamos una (reunión) de cierta k -muestra de los conjuntos:

$$S_{i_1} \cup S_{i_2} \cup \dots \cup S_{i_k};$$

se deben examinar dos casos: uno cuando el número de elementos en la unión es igual a k , y otro cuando el número de elementos es superior a k . Empecemos por el segundo caso, cuando la unión en consideración posee no menos de $k+1$ elementos, cualesquiera que sean k : $k = 1, 2, \dots, m$, y el juego $i_1, i_2, \dots, i_k$ de números $1, 2, \dots, m$. Elijamos algún elemento $a_1 \in S_{i_1}$ y eliminémoslo en $S_2, S_3, \dots, S_m$, si él se encuentra en dichos conjuntos. Llegamos de este modo a $M^* = (S_2^*, S_3^*, \dots, S_m^*)$. Esta $(m-1)$ -muestra satisface la condición necesaria para la existencia de un s.r.d., puesto que $S_{i_1}^* \cup S_{i_2}^* \cup \dots \cup S_{i_k}^*$ contiene no menos de k elementos y, además, M^* tiene o bien no menos de $(t-1)$ s.r.d., si $t \leq m$ (y, por tanto, $t-1 \leq m-1$), o bien $(t-1)!/(t-m)!$ s.r.d., si $t > m$ (y, por tanto $t-1 > m-1$). El resultado buscado se logra, si tomamos en consideración que en S_1 hay por lo menos t posibilidades de fijar el elemento a_1 y que este elemento constituye, junto con el s.r.d. para M^* , el s.r.d. para M .

Volvamos al primer caso en que la demostración realizada no puede

considerarse válida, pues existe una k -muestra $S_{i_1}, S_{i_2}, \dots, S_{i_k}$ tal que $S_{i_1} \cup S_{i_2} \cup \dots \cup S_{i_k}$ contiene exactamente k elementos ($1 \leq k \leq m-1$). Reenumeremos los conjuntos $S_1, S_2, \dots, S_m$ de un modo tal que S_{i_1} se haga S_1 , S_{i_2} se haga S_2 , $\dots$, y S_{i_k} se haga S_k ; escribamos estos conjuntos en el siguiente orden nuevo:

$$S_1, S_2, \dots, S_k, S_{k+1}, \dots, S_m.$$

Por cuanto $S_1 \cup S_2 \cup \dots \cup S_k$ contiene exactamente k elementos, entonces $t \leq k$. Por consiguiente, por hipótesis de inducción, $(S_1, S_2, \dots, S_k)$ tiene no menos de $t!$ s.r.d. Tomemos uno de estos s.r.d.: $(a_1, a_2, \dots, a_k)$, donde $a_i \in S_i$, $i = 1, 2, \dots, k$. Eliminemos los elementos $a_1, a_2, \dots, a_k$ de $S_{k+1}, \dots, S_m$, siempre que se encuentren allí. La $(m-k)$ -muestra obtenida $M^* = (S_{k+1}^*, \dots, S_m^*)$ satisface la condición necesaria para que existe un s.r.d. Efectivamente, si esto no es así y si la unión de cierta k^* -muestra $S_{k+1}^* \cup \dots \cup S_{k+k^*}^*$ tiene menos de k^* elementos, entonces la unión

$$S_1 \cup S_2 \cup \dots \cup S_k \cup S_{k+1}^* \cup \dots \cup S_{k+k^*}^*$$

tendría menos de $k + k^*$ elementos, lo que contradice la hipótesis del teorema. Así pues, M^* tiene por lo menos un s.r.d., y, por consiguiente, M tiene no menos de $t!$ s.r.d.

Algoritmo de elección de un s.r.d. Prácticamente es muy difícil comprobar si en este caso concreto se cumple o no la hipótesis del teorema de P. Hall. La demostración que acabamos de ofrecer, basada en la inducción matemática completa, no proporciona ninguna indicación que ayude a hallar el s.r.d. Esto no es sorprendente. Los teoremas de existencia aparecen, las más de las veces, cuando resulta difícil o imposible hallar un algoritmo que conduzca a la determinación de la solución. El algoritmo que permite elegir un s.r.d. para un número finito de conjuntos, o mostrar que para el juego dado de conjuntos tal sistema no existe, lo dió M. Hall como demostración del teorema de P. Hall sobre los distintos representantes.

Sean dados n conjuntos $S_1, S_2, \dots, S_n$. Se requiere hallar para ellos el s.r.d., o mostrar que tal sistema no existe. Elijamos al azar un elemento del primer conjunto $a_1 \in S_1$ en calidad de representante de este último. Elegiremos por turno los representantes de otros conjuntos: $a_2 \in S_2$; $a_3 \in S_3$; $\dots$, preocupándonos sólo de que todos estos representantes sean distintos. Si llevamos este proceso hasta $a_n \in S_n$ inclusive, obtendremos el s.r.d. buscado.

Puede ocurrir que en un r -ésimo paso llegaremos a conseguir cierto t -conjunto S_r , todos los elementos del cual $b_1, b_2, \dots, b_t$ ya habían sido elegidos en calidad de representantes de otros conjuntos. Esto, sin embargo, no significa todavía que el s.r.d. no existe. Vamos a tomar, uno tras otro, todos aquellos conjuntos, cuyos representantes son los elementos b_i ($i = 1, 2, \dots$), y eliminar de ellos todos los elementos de la sucesión $b_1, b_2, \dots, b_t$, agregando los elementos restantes al final de la sucesión citada. Procedamos de este modo hasta que ocurra una cosa de dos: o bien 1) llegaremos al elemen-

to b_i , el cual no puede servir de representante, o bien 2) la sucesión se agotará con los elementos $b_1, b_2, \dots, b_s$ como representantes de los conjuntos.

En el caso 2) podemos estar convencidos de que el s.r.d. no existe. En efecto, los elementos $b_1, b_2, \dots, b_s$ son representantes de s conjuntos y, por construcción, cada elemento de estos s conjuntos se contiene en la sucesión dada. Pero, en tal caso, los s conjuntos mencionados, y también el conjunto S_r , forman $s + 1$ conjuntos que sólo contienen s elementos distintos, lo que contradice la hipótesis del teorema.

En cambio, si tiene lugar el caso 1), entonces encontramos, en cierta etapa, el elemento $b_i = b_{i_1} \in S_{j_1}$ ($i_1 > i$) que no ha sido hasta ahora un representante. Esto es testimonio de que como representante de S_{j_1} ya fue elegido otro elemento b_{i_2} ($i_2 < i_1$). Si $i_2 > i$, entonces el elemento b_{i_2} pertenece al conjunto S_{j_2} cuyo representante es b_{i_3} ($i_3 < i_2$), etc. Surge, pues, una sucesión $b_{i_1}, b_{i_2}, \dots, b_{i_m}$, cuyos índices decrecen ($i_m \leq i$), con la particularidad de que cada elemento de esta sucesión integra un conjunto, cuyo representante es el término sucesivo. Sustituyamos los representantes, eligiendo los elementos del modo siguiente: b_{i_1} para S_{j_2} , b_{i_2} para S_{j_3} , $\dots$, $b_{i_{m-1}}$ para $S_{j_{m+1}}$. Como resultado de esta sustitución, el elemento b_{i_m} se libera para poder ser elegido en calidad de representante de S_r . Así pues, $S_1, \dots, S_r$ tienen representantes distintos y podemos seguir la misma senda, teniendo en cuenta o bien la posibilidad de llegar a S_n y obtener el s.r.d. completo, o bien encontrarse con el caso 2) y establecer que el s.r.d. no existe.

La conclusión sobre el número de s.r.d. se obtiene del algoritmo mencionado como una consecuencia. Efectivamente, si el s.r.d. existe, esto significa que existe también un conjunto, cada elemento del cual puede ser elegido en calidad de su representante en el s.r.d. Esto quiere decir, que si los conjuntos de una familia dada de n elementos tienen t o más elementos, entonces existen por lo menos $t!$ s.r.d., si $t < n$, o bien $t(t-1) \dots (t-n+1)$ elementos, si $t \geq n$ (por cuanto la elección del primer representante puede realizarse empleando por lo menos t métodos; al tachar este representante elegido por nosotros en todos los demás conjuntos, llegaremos al sistema de conjuntos $S_2^*, \dots, S_n^*$, donde el conjunto menor tiene no menos de $t-1$ elementos. Continuando así, paso a paso, obtendremos el resultado mencionado).

La condición de que el sistema sea finito es aquí de mucha importancia. Si retiramos esta condición, entonces, para, por ejemplo, un sistema infinito de conjuntos

$$S_0 = \{1, 2, 3, \dots, k, \dots\};$$

$$S_1 = \{1\};$$

$$\dots \dots \dots$$

$$S_k = \{k\};$$

$$\dots \dots \dots$$

no existe un s.r.d., aunque sí existe para cualquier parte de él.

Otros sistemas de representantes de los conjuntos. Los problemas de partición de los conjuntos condujeron al concepto de sistemas de representantes comunes. Sean dadas dos particiones diferentes de un mismo conjunto S en k componentes no vacías:

$$S = A_1 \cup A_2 \cup \dots \cup A_k; \quad S = B_1 \cup B_2 \cup \dots \cup B_k.$$

Si existe un subconjunto O del conjunto S , compuesto de k elementos, y si, además, dicho subconjunto es tal que su intersección con cualquiera de las componentes no es vacía:

$$O \cap A_i \neq \emptyset; \quad O \cap B_i \neq \emptyset; \quad i = 1, 2, \dots, k,$$

éste se denominará *sistema de representantes comunes* (s.r.c.) de las particiones dadas. En este caso cada una de las intersecciones resulta estar compuesta por un solo elemento. Los conjuntos de las particiones primera y segunda, tomados dos a dos y elegidos, si es necesario, de un modo correspondiente, tienen uno y sólo un elemento común, el cual es precisamente su representante común.

No es forzosamente obligatorio, por supuesto, que se cumpla el requerimiento de existencia de un solo elemento común. Pueden plantearse y tomarse en consideración condiciones diferentes: existencia de un número dado de elementos comunes, de un conjunto dado, etc. Por ejemplo, si exigimos que cada elemento $s_i \in S$ figure no menos de k_{1i} y no más de k_{2i} veces ($k_{2i} \geq k_{1i} \geq 1$) en el sistema de representantes comunes, este último se llamará sistema de un número limitado de representantes [28]. Está claro que el problema de s.r.d es un caso particular de tal problema para $k_{1i} = 0$, $k_{2i} = 1$.

El otro caso particular en que

$$k_{1i} = \begin{cases} 1 & \text{para } i = 1, 2, \dots, l; \\ 0 & \text{para } i = l+1, l+2, \dots, m; \end{cases}$$

$$k_{2i} = 1 \text{ para } i = 1, 2, \dots, m,$$

lleva el nombre de problema sobre la existencia del s.r.d. que contiene el conjunto dado de elementos marginales $s_1, s_2, \dots, s_l$.

El criterio de existencia o no existencia del s.r.c. es próximo al criterio que se aplica al s.r.d.

Teorema. Dos particiones de un conjunto

$$S = A_1 \cup A_2 \cup \dots \cup A_k = B_1 \cup B_2 \cup \dots \cup B_k$$

tienen un s.r.d., cuando y sólo cuando la unión de cualesquiera m del conjunto A_i se intersecan no menos que con m del conjunto B_j , donde $m = 1, \dots, k$.

Demostración. La necesidad es obvia, igual que en el caso del s.r.d. La suficiencia se demuestra por reducción a los teoremas sobre el s.r.d.

En efecto, elijamos para cada B_i , $i = 1, 2, \dots, k$, un conjunto S_i de todos los índices $j \in K = \{1, \dots, k\}$ tales que $A_j \cap B_i \neq \emptyset$. Obtendremos una m -muestra $M = (S_1, S_2, \dots, S_k)$ de subconjuntos del conjunto K . Para M

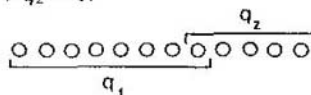
existe el s.r.d. (el criterio enunciado de existencia de un s.r.d. es el criterio de existencia del s.r.d. para M). La elección de los distintos representantes da para cada B_i su A_j , con la particularidad de que su intersección es no vacía. En esta intersección puede elegirse por lo menos un elemento que sea común para A_j y B_i , es decir, su representante común.

El concepto de representantes de los conjuntos y de sistemas de representantes tiene en las matemáticas numerosas y variadas aplicaciones, como son, por ejemplo, los representantes de las clases de equivalencia. El método de sistemas de representantes se usa también en la teoría de las redes, al investigar la admisibilidad de los flujos, y en la teoría de ampliación de los cuadrados latinos.

3.3. PRINCIPIOS DE LA TEORÍA DE RAMSEY

La operación de partición de los conjuntos es el fundamento de una cantidad infinita de problemas prácticos. En este párrafo se describe el procedimiento que permite obtener los datos sobre el carácter de las propias particiones y sobre la posibilidad de realizar una partición del tipo prefijado de antemano.

Sea un n -conjunto S , cuyos elementos han de distribuirse en dos cajones. ¿Cuál debe ser el número n , para asegurar que q_1 elementos caigan en el primer cajón, o que q_2 elementos caigan en el segundo? La respuesta es casi evidente: $n \geq q_1 + q_2 - 1$.



Así pues, el número mínimo de elementos que aseguran la resolución del problema es igual a

$$N(q_1, q_2; 1) = q_1 + q_2 - 1.$$

Si se plantea un problema de distribución no entre dos, sino entre el mayor número (digamos, t) de cajones, siendo $q_1, q_2, \dots, q_t$ los números correspondientes que caracterizan el llenado requerido de los cajones, entonces

$$N(q_1, q_2, \dots, q_t; 1) = \sum_{i=1}^t q_i - (t - 1).$$

Ramsey [45] demostró en 1930 un teorema en el que generalizó estos resultados particulares. En lugar de la distribución de los elementos unidad del n -conjunto el teorema analiza la distribución de los r -subconjuntos del conjunto citado. Designemos con S_n el conjunto de n elementos y con $P_r(S_n)$, la colección de sus r -subconjuntos.

Teorema de Ramsey. Supongamos que $r \geq 1$, $q_i \geq r$ ($i = 1, \dots, t$). Existe un número natural mínimo $N = N(q_1, \dots, q_t; r)$ tal que para cualquier

$n \geq N$ y para toda t -partición ordenada $P_r(S_n) = A_1 \cup \dots \cup A_t$ se encontrará (para cierto $i \in \{1, \dots, t\}$) un (q_i, A_i) -subconjunto, es decir, un q_i -subconjunto del conjunto S_n , cuyos r -subconjuntos están contenidos todos en A_i .

Hagamos previamente tres observaciones con el fin de facilitar la comprensión del teorema en su planteamiento y demostración.

1. Los ejemplos aducidos más arriba y referentes a las particiones de los conjuntos representan casos particulares del teorema de Ramsey para $r = 1$, $P_r(S) = S$, y el (q_i, A_i) -conjunto es simplemente un q_i -subconjunto del conjunto A_i .

2. Para un caso particular en que $t = 1$, tenemos $N(q_1, r) = q_1$, lo que es trivial.

3. Si demostramos el teorema de Ramsey para $t = 2$, quedará válido también para $t = 3$. En efecto, sea

$$P_r(S) = A_1 \cup A_2 \cup A_3 = A_1 \cup (A_2 \cup A_3).$$

Denotemos $A_2 \cup A_3 = A_2^*$. Pongamos $q_2^* = N(q_2, q_3; r)$. Si $n \geq N(q_1, q_2^*; r)$, entonces, o bien S contiene el (q_1, A_1) -subconjunto, o bien S contiene el $(q_2^*, A_2 \cup A_3)$ -subconjunto. El primer caso corresponde a la afirmación del teorema. En cambio, si tiene lugar el segundo caso, entonces el q_2^* -subconjunto del conjunto S contendrá, por hipótesis, o bien el (q_2, A_2) -conjunto o bien el (q_3, A_3) -subconjunto. Por lo tanto, nuestra afirmación es lícita. De aquí se deduce que resulta suficiente demostrar el teorema de Ramsey para $t = 2$, y, a continuación concluir por inducción, que es válido también para todo $t > 2$ entero.

La demostración del teorema para el caso de $t = 2$ empieza escribiendo las siguientes igualdades, considerándolas como iniciales:

$$N(q_1, q_2; 1) = q_1 + q_2 - 1; \quad N(q_1, r; r) = q_1; \quad N(r, q_2; r) = q_2. \quad (1)$$

La primera de estas igualdades ya la conocemos. En la segunda cualquier partición $A_1 \cup A_2$ lleva o bien al (r, A_2) -subconjunto, si A_2 es no vacío, o bien al $(q_1, P_r(S))$ -subconjunto. En este caso q_1 no es inferior a r . Cualquiera de estos casos corresponde exactamente a la formulación del teorema. Unos razonamientos no complejos mostrarán la validez también de la tercera igualdad. Además,

$$r > 1; \quad q_1 \geq r; \quad q_2 \geq r. \quad (2)$$

Así pues, ya podemos afirmar que existen los números $N(2, 2, 2)$, $N(3, 2, 2)$, $N(4, 2, 2)$ etc., como también $N(2, 3, 2)$, $N(2, 4, 2)$, etc. Las igualdades (1) y las desigualdades (2) sirven de datos iniciales para la demostración del teorema de Ramsey por inducción. El primer paso, pues, lo hicimos. El segundo paso, que da por terminada la demostración, consiste en demostrar la existencia de $N(q_1, q_2; r)$ a condición de que existen

$$p_1 = N(q_1 - 1, q_2; r); \quad p_2 = N(q_1, q_2 - 1; r); \quad N(p_1, p_2; r - 1).$$

Por cuanto hasta el momento no se logra calcular exactamente $N(q_1, q_2; r)$, el sentido de la demostración consiste en estimar este número superiormente, demostrando la desigualdad

$$N(q_1, q_2; r) \leq N(p_1, p_2; r-1) + 1.$$

En efecto, tomemos un conjunto S_n , donde $n \geq N(p_1, p_2; r-1) + 1$, y fijemos en él un elemento a_0 ; entonces, $S_n \setminus \{a_0\} = S_{n-1}$. Pasemos de la partición $P_r(S_n) = A_1 \cup A_2$ a la $P_{r-1}(S_{n-1}) = B_1 \cup B_2$ del modo siguiente: en el conjunto S_{n-1} tomamos por turno los $(r-1)$ -subconjuntos; si la unión de tal subconjunto, digamos, de B' , y del elemento fijo a_0 integra A_1 , entonces este $(r-1)$ -subconjunto B' lo atribuiremos a la clase B_1 , y si la unión está contenida en A_2 , entonces B' lo atribuiremos a la clase B_2 .

El conjunto S_{n-1} tiene no menos de $N(p_1, p_2; r-1)$ elementos y para este conjunto el teorema es válido por hipótesis de inducción. Por consiguiente, en él se contiene o bien a) el (p_1, B_1) -subconjunto, o bien b) el (p_2, B_2) -subconjunto. Supongamos que tiene lugar el caso a). Esto significa que en S_{n-1} se contiene el p_1 -conjunto T , cuyos $(r-1)$ -subconjuntos están contenidos todos en B_1 . Pero, $p_1 = N(q_1-1, q_2; r)$; $T \subseteq S_n$, y en este conjunto T existe o bien un (q_1-1) -subconjunto, cuyos r -subconjuntos integran todos A_1 , o bien un q_2 -subconjunto, cuyos r -subconjuntos integran todos A_2 . Si tiene lugar el último caso, entonces obtendremos el q_2 -conjunto que satisface las condiciones del teorema. Si en cambio, tiene lugar el primer caso, es decir, existe el (q_1-1) -subconjunto, cuyos r -subconjuntos integran todos A_1 , entonces reunámoslo con el elemento fijo a_0 . Obtendremos el q_1 -conjunto integrado en S . Veamos sus r -subconjuntos. Si alguno de ellos no contiene a_0 , será el r -conjunto que está contenido en A_1 . Si él contiene a_0 , entonces está compuesto por este elemento y el $(r-1)$ -subconjunto que integra B_1 . Esto quiere decir que está integrado en A_1 . Hemos obtenido el q_1 -subconjunto del conjunto S_n , todos los r -subconjuntos del cual están contenidos en A_1 .

Se ha examinado el caso en que el (p_1, B_1) -conjunto está contenido en S_{n-1} . El caso b) se demuestra de una manera análoga. El teorema de Ramsey queda demostrado.

A lo que parece para toda t -partición y para cualesquiera $(q_1, q_2, \dots, q_t; r)$ convenientes, se puede determinar mín $n = N(q_1, q_2, \dots, q_t; r)$, cuya existencia está asegurada en virtud del teorema de Ramsey. Mas, este trabajo resultó ser extremadamente difícil, pues por ahora no se conoce ningún método para calcular estos números. Para $t = r = 2$, todos los números conocidos $N(q_1, q_2; 2)$ se aducen en la Tabla 3.1 (si no se conocen los valores exactos, se dan las fronteras dentro de las cuales éstos están encerrados):

Los casos en que $t > 2$ están poco estudiados por ahora. Se sabe, por ejemplo, que

$$\begin{aligned} N(3, 3, 3; 2) &= 17; & N(5, 5, 5; 2) &\geq 257; \\ N(4, 4, 4; 2) &\geq 128 & N(6, 6, 6; 2) &\geq 906. \end{aligned}$$

Tabla 3.1

	3	4	5	6	7	8	9
3	6	9	14	18	23	28 ... 29	36
4	9	18	25 ... 28	34 ... 44			
5	14	25 ... 28	42 ... 55	57 ... 94			
6	18	34 ... 44	57 ... 94	102 ... 169			
7	23						
8	28 ... 29						
9	36						

Aún menos sabemos sobre los valores de los números de Ramsey para $r > 2$. Merece la atención el siguiente resultado importante:

$$13 \leq N(4, 4; 3) \leq 15.$$

Pese a las excepcionales dificultades y al avance lento, no cesan los esfuerzos que tienen por objeto determinar los números de Ramsey o de sus estimaciones (superiores e inferiores). Esto se explica por la peculiaridad del problema que admite numerosas interpretaciones y aplicaciones diversas. En la teoría de los grafos, por ejemplo, este problema se interpreta como un problema de coloración de las aristas de los grafos. El propio problema de buscar los números de Ramsey se considera, a menudo, para las clases aisladas de grafos.

Otra dirección, en la que se desarrolla la teoría de Ramsey, representa el estudio de las condiciones de existencia, es decir, de tales restricciones, bajo las cuales los planteamientos generalizados del problema resulten fundamentados, y las demostraciones, factibles.

Los problemas de la teoría de Ramsey se estudian ahora con un elevado grado de generalidad. Supongamos que el símbolo $r_k \nabla (l_1, \dots, l_r)$ significa lo siguiente: si los k -conjuntos de un n -conjunto S_n están partidos en r clases, entonces para cierto i existe un l_i -subconjunto $L_i \subseteq S_n$ tal que todos los k -subconjuntos de L_i integran la i -ésima clase. En este caso el teorema de Ramsey se enuncia del modo siguiente:

Teorema. Para cualesquiera $k, r, l_1, \dots, l_r$ naturales existe un $N = N(k, r, l_1, l_r)$ tal que si $n \geq N$, entonces $r_k \nabla (l_1, \dots, l_r)$.

En general, los teoremas de tipo Ramsey tienen la forma de las afirmaciones $r_k \nabla (C_1, \dots, C_r)$, donde los símbolos $A, B, C_1, \dots, C_r$ designan objetos (del conjunto) de estructura determinada. Los últimos pueden ser conjuntos con relaciones de orden sobre ellos, grafos e hipergrafos, espacios vectoriales finitos, conjuntos de soluciones de los sistemas de ecuaciones lineales, álgebras de Bool y particiones de los conjuntos finitos. Las generalizaciones relacionadas con el caso de los conjuntos finitos conducen al análisis de las operaciones con números cardinales. Los resultados conseguidos en este camino por ahora no son grandes.

CAPÍTULO 4 APARATO TABULAR DE MATRIZ DEL ANÁLISIS COMBINATORIO

En este capítulo se analizan los conceptos y métodos combinatorios que están relacionados con la representación de los sistemas de conjuntos finitos en forma de tablas. Sirve de base para tales representaciones tabulares el concepto general de sistemas de incidencia.

4.1. SISTEMAS DE INCIDENCIA Y MATRICES ESPECIALES

Se denomina *sistema de incidencia* una terna ordenada (M, S, φ) , donde φ es una correspondencia binaria entre los conjuntos M y S (véase § 1.2). Por ejemplo, un plano puede ser representado como sistema de incidencia, si tomamos en calidad de S el conjunto de puntos del plano, en calidad de M , el conjunto de rectas y en calidad de φ , el conjunto de pares $(a, b) \in M \times S$ que satisfacen la condición "el punto b se dispone en la recta a ". La correspondencia binaria φ en el sistema de incidencia recibe el nombre de *relación de incidencia*. En la mayoría de los casos los elementos del conjunto M se interpretan como ciertos subconjuntos del conjunto S , mientras que φ es simplemente una relación que caracteriza la pertenencia de los elementos del conjunto S a los subconjuntos de la familia M . Si $M = \{S_1, \dots, S_m\}$, $S = \{s_1, \dots, s_n\}$, entonces a la relación φ le corresponderá biunívocamente la matriz $A = \|a_{ij}\|$ de orden $m \times n$, donde a_{ij} (elemento de la i -ésima fila y de la j -ésima columna) es igual a la unidad, siempre que $(S_i, s_j) \in \varphi$, y a cero, en el caso contrario. En particular, cuando $M = \{S_1, \dots, S_m\}$ se interpreta como un sistema de subconjuntos del conjunto S , entonces

$$a_{ij} = \begin{cases} 1, & \text{si } s_j \in S_i; \\ 0, & \text{si } s_j \notin S_i. \end{cases}$$

Esta matriz se llama *matriz de incidencia* para la familia de subconjuntos $M = \{S_1, \dots, S_m\}$ respecto del conjunto S . Las unidades dispuestas en la i -ésima fila expresan elementos del subconjunto S_i , y las unidades dispuestas en la j -ésima columna señalan los subconjuntos que contienen s_j . La matriz A da una descripción completa del sistema de incidencia (M, S, φ) . Es *binaria*, esto es, se compone de símbolos de los tipos. Para que los cálculos sean más cómodos, se ha elegido aquí, a título de elementos de la matriz, 0 y 1 (en adelante las matrices de esta índole se llamarán $(0, 1)$ -matrices), aunque se eligen, a veces, otros pares de magnitudes, por ejemplo, $+1$ y -1 .

El dominio de aplicación de las matrices de incidencia es muy amplio. Kirchhoff, por ejemplo, las introdujo para investigar los circuitos eléctricos; A. Poincaré, en la topología, etc.

Los teoremas del análisis combinatorio se interpretan frecuentemente e, incluso, se demuestran con ayuda de las matrices de incidencia. Por ejemplo, en el § 3.2 se ha enunciado el teorema de P. Hall sobre el sistema de representantes distintos. De su análogo matricial sirve el

Teorema de König [46]. En una matriz rectangular el número mínimo m de líneas (de filas y de columnas) que contienen todos los elementos no nulos, es igual al número máximo M de dichos elementos elegidos de un modo tal que de ellos cualesquiera dos no están dispuestos en una línea.

Demostración. Basta estudiar el caso en que viene dada una $(0,1)$ -matriz rectangular $A = [a_{ij}]$, $i = 1, 2, \dots, n$; $j = 1, 2, \dots, l$. La demostración de que $m = M$ consta de dos etapas; al principio se demuestra que $m \geq M$, luego que $M \geq m$. La primera desigualdad es obvia, pues ninguna línea contiene más de un elemento de los M elegidos. Demostremos la segunda desigualdad.

Supongamos que m líneas están compuestas por r filas y s columnas: $m = r + s$. Permutemos estas filas y columnas de un modo tal que ellas ocupen los primeros lugares. A cada j -ésima fila ($i = 1, \dots, r$) pongámonle en correspondencia un conjunto de números de las columnas j , para las cuales $a_{ij} = 1$, $j > s$. Los conjuntos obtenidos satisfacen las condiciones del teorema de P. Hall. Efectivamente, si esto no fuera así, es decir, si k de los conjuntos citados contienen $v < k$ elementos, dichas k filas pueden ser sustituidas por v columnas y todas las unidades figurarán en el número menor de líneas, lo que contradice la condición de que m es el número mínimo. Por cuanto la condición del teorema de P. Hall se cumple, se pueden elegir r representantes distintos de r filas, es decir, r unidades de manera tal que haya dos unidades en una fila y ninguna, en las primeras s columnas. Razonando análogamente, se pueden obtener s representantes de s primeras columnas de un modo tal que estén ausentes en las primeras r filas. Este conjunto de $r + s = m$ unidades está elegido de una manera tal que no hay dos de ellas que se dispongan en una línea. Esto quiere decir, que $m \leq M$, lo que se trataba de demostrar.

La demostración aludida aquí se apoya en el teorema de P. Hall. Este último puede demostrarse también apoyándose en el teorema de König. Sean un conjunto S y un sistema de sus n subconjuntos. Construyamos la $(0, 1)$ -matriz de incidencia. Supongamos que todas las unidades de ésta se disponen en r filas y s columnas. Si $r + s = n$, entonces, según el teorema de König, existen n unidades, de las cuales cualesquiera dos no están en una misma línea, y ellas forman el s.r.d para n conjuntos. Si, en cambio, $r + s < n$, entonces la validez del teorema de P. Hall se perturba, puesto que para $k = n - r$ filas las unidades figurarán sólo en $s < n - r = k$ columnas.

Las matrices binarias se utilizan en el análisis combinatorio al investigar las estructuras de los conjuntos discretos. La teoría de las matrices binarias experimentó en los últimos años un desarrollo considerable precisamente merced al papel creciente (práctico y teórico) del análisis combinatorio.

Según sea el tipo de problemas combinatorios, en el proceso de su resolución aparecen diferentes tipos de matrices. Las operaciones sobre ellas se realizan, por supuesto, de acuerdo con las reglas generales de la teoría de matrices. Así, por ejemplo, a la par con las sumas y los productos ordinarios, están determinadas para ellas las sumas y los productos directos (de Kronecker). Recordemos que la *suma directa* de la matriz $A = [a_{ij}]$ de n -ésimo orden con la matriz B de orden m se representa por una matriz de orden $n + m$ que tiene la forma

$$\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix},$$

y el *producto directo* de estas matrices es una matriz de orden nm que tiene la forma

$$\begin{pmatrix} a_{11}B & \dots & a_{1n}B \\ \dots & \dots & \dots \\ a_{n1}B & \dots & a_{nn}B \end{pmatrix}.$$

Más abajo estudiaremos ciertos tipos especiales de las matrices que con mayor frecuencia se usan en el análisis combinatorio.

Matrices conmutables. Son matrices de incidencia para el caso en que $S = M = \{1, \dots, n\}$, y la relación de incidencia ij significa que el elemento i se dispone en el j -ésimo lugar de cierta permutación fija de números $1, \dots, n$. Sea, por ejemplo, una permutación $(5, 3, 4, 1, 2)$. La matriz correspondiente será

$$\begin{pmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Sus filas corresponden a los números de los elementos de la permutación, y las columnas, a los números ordinales de los lugares.

Así pues, una matriz conmutable es la $(0, 1)$ -matriz cuadrada, en la cual cada línea (fila o columna) contiene exactamente una unidad. La matriz conmutable posee la propiedad de que

$$PP^T = I,$$

donde P^T es una matriz transpuesta, e I , una matriz unidad. Esta propiedad se emplea con frecuencia a título de definición de la matriz conmutable.

Las matrices conmutables, siendo las más simples, son de interés en primer lugar por el hecho de que a través de ellas pueden expresarse otras matrices más complejas (véase más abajo el teorema de Birkhoff).

Matrices de congruencia de dos en dos. Estas son matrices de incidencia para el caso en que $M = S$, y la relación de incidencia φ es antisimétrica y satisface las condiciones:

- a) para cualquier $s \in S$: $(s, s) \notin \varphi$ (irreflexividad);
- b) para cualesquiera diferentes $s_i, s_j \in S$ ($i \neq j$): o bien $(s_i, s_j) \in \varphi$, o bien $(s_j, s_i) \in \varphi$.

Tal relación puede interpretarse como resultado de un torneo circular con participación de $|S|$ jugadores. La matriz de incidencia es la tabla de este torneo, en la que la diagonal está ocupada por ceros ($a_{ii} = 0$). El premio del i -ésimo jugador y del j -ésimo se nota como $a_{ij} = 1$, $a_{ji} = 0$.

Así pues, las matrices de congruencia de dos en dos son binarias, cuadradas, con ceros en la diagonal principal. Ellas satisfacen la ecuación matricial

$$A + A^T = J - I,$$

donde A es una matriz de congruencia de dos en dos, A^T es la matriz transpuesta A , I es una matriz unidad y J , una matriz compuesta por ceros.

Matrices de Hadamard. Así se llaman las matrices cuadradas binarias H (abreviadamente de Hadamard), compuestas de unidades positivas y negativas, que satisfacen la ecuación matricial

$$H \cdot H^T = nI,$$

donde n es el orden de la matriz H ; I es la matriz unidad de n -ésimo orden: H^T , la matriz H transpuesta. De la definición proviene que cada dos filas de la matriz H son ortogonales, y, además, que $|\det H| = n^{n/2}$ y $H^{-1} = n^{-1}H^T$, de donde

$$H \cdot H^T = H^T \cdot H = nI.$$

Si en una matriz de Hadamard multiplicamos cualquier línea (sea fila o columna) por -1 , la matriz conserva sus cualidades de Hadamard. Las matrices obtenidas del modo citado se consideran equivalentes. Si H_1 y H_2 son equivalentes, entonces

$$H_2 = PH_1Q,$$

donde P y Q son matrices conmutables, en las cuales, sin embargo, el único elemento no nulo de la línea es igual a $+1$, o bien a -1 . Multiplicando las filas o columnas por -1 , podemos normalizar las matrices de Hadamard, es decir, reducirlas a una forma en que la primera fila y la primera columna están compuestas sólo por las unidades positivas. Las matrices de Hadamard normalizadas de órdenes 1 y 2 son

$$(1) \quad \text{y} \quad \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

En lo que atañe a las matrices de orden $n \geq 3$, resulta que si existen, entonces $n \equiv 0 \pmod{4}$. Efectivamente, conmutemos en la matriz de

Hadamard las columnas de un modo tal que en la segunda fila la primera mitad de términos conste sólo de unidades positivas, y la segunda mitad, de unidades negativas. Examinemos la tercera fila. Supongamos que en los primeros $n/2$ lugares encontramos t unidades positivas y en los segundos t' unidades positivas. Entonces, por definición,

$$2t + 2t' = n;$$

$$2t - 2t' = 0,$$

de donde $n = 4t$. No obstante, la demostración de la existencia de las matrices en el sentido de su construcción eficaz es un proceso que avanza con lentitud, pese a la diversidad de los métodos inventados.

Observemos que los productos directos de las matrices de Hadamard son también matrices de Hadamard. Se sabe que estas matrices de orden n existen, en particular, en los siguientes casos (véase [47]):

1. $n = 2^r$;
2. $n = p^r + 1 \equiv 0 \pmod{4}$, donde p es un número primo;
3. $n = (h - 1)^3 + 1$, donde h es un producto de los números del tipo 1 y 2;
4. $n = h(h + 3)$, donde h y $h + 4$ son productos de los números del tipo 1 y 2;
5. $n = h(h - 1)$, donde h es un producto de los números del tipo 1 y 2;
6. $n = 92, 116, 156, 172$;
7. $n = q(q + 2) + 1$, donde q y $q + 2$ son potencias de los números primos;
8. n es igual al producto de cualesquiera números mencionados más arriba.

Para el caso incierto mínimo de existencia de las matrices de Hadamard tenemos: $n = 268$.

Matrices estocásticas. Este tipo de matrices surgió para ser aplicado en el aparato combinatorio de la teoría de probabilidades. Una matriz se llama estocástica por filas (columnas), si sus elementos son no negativos y la suma de ellos en cada fila (columna) es igual a 1. Una matriz se llama dos veces estocástica, si es estocástica tanto por filas, como por columnas. No es difícil mostrar que una matriz dos veces estocástica es cuadrada. Un producto de las matrices dos veces estocásticas de igual orden será matriz dos veces estocástica. Efectivamente, si $C = [c_{ij}]_n^1$, $A = [a_{ij}]_n^1$, $B = [b_{ij}]_n^1$, $C = A \cdot B$, entonces

$$c_{ij} = \sum_{l=1}^n a_{il} b_{lj} \geq 0, \quad \sum_{j=1}^n c_{ij} = \sum_{l=1}^n a_{il} \sum_{j=1}^n b_{lj} = 1,$$

y, por analogía, $\sum_{i=1}^n c_{ij} = 1$.

En una matriz cuadrada de n -ésimo orden, se llamará *transversal* a un juego de n elementos, en el que ningún par de ellos está dispuesto en una misma línea (fila o columna).

Lema. Una matriz dos veces estocástica tiene la transversal compuesta de los elementos no nulos.

Demostración. Sea $A = [a_{ij}]$ una matriz dos veces estocástica. Incluyamos todos los elementos no nulos de la matriz en p columnas y q filas de un modo tal que la suma $p + q$ asuma el valor mínimo posible. Con arreglo al teorema de König, el número máximo de elementos no nulos de A , de los cuales ningún par se dispone en una misma línea, es igual a $p + q$. Por cuanto todos los elementos no nulos de A pueden encerrarse dentro de n filas, entonces $p + q \leq n$. Por otra parte, $n = \sum_{i,j=1}^n a_{ij} \leq p + q$. Por consiguiente, $p + q = n$, y en la matriz se en-

contrará una transversal compuesta por elementos no nulos.

Teorema de Birkhoff. Para toda matriz dos veces estocástica de orden n es válido el desarrollo: $A = \sum_{i=1}^k \lambda_i P_i$, donde P_i son matrices conmutables, $\lambda_i > 0$, $\sum_{i=1}^k \lambda_i = 1$, $k \leq n^2 - n + 1$.

Demostración. De acuerdo con el lema, en la matriz A existe una transversal compuesta por elementos no nulos. Supongamos que el mínimo de dichos elementos es igual a $\lambda_1 > 0$ ($\lambda_1 \neq 1$, pues, de lo contrario, la matriz A será conmutable y la afirmación del teorema se hará trivial): P_1 es una matriz conmutable correspondiente a la transversal. Por cuanto la matriz $(1 - \lambda_1)^{-1}(A - \lambda_1 P_1)$ es dos veces estocástica, entonces en la matriz $A - \lambda_1 P_1$ existe una transversal compuesta de los elementos no nulos. Designemos con λ_2 el mínimo de ellos, etc. La matriz $A - \sum_{i=1}^j \lambda_i P_i$ contiene no menos de j elementos nulos, razón por la cual el proceso citado es finito y se dará por terminado no más tarde que el $(n^2 - n + 1)$ -ésimo paso. Esto se debe a que después de realizarse el $(n^2 - n + 1)$ -ésimo paso, la matriz $A - \sum_{i=1}^{n^2-n+1} \lambda_i P_i$, en la que las sumas de elementos en las filas y en las columnas son iguales, contendría no más de $(n - 1)$ elementos no nulos, por lo cual todos los elementos de esta matriz serían nulos.

Observemos que la estimación del número k en el teorema de Birkhoff puede ser precisada: $k \leq n^2 - 2n + 2$, con la particularidad de que existen matrices de orden n , dos veces estocásticas, las cuales no son representables en forma de una combinación lineal de menos que $n^2 - 2n + 2$ matrices conmutables.

Al resolver problemas de carácter combinatorio surge la necesidad de emplear, además, otros tipos de matrices especiales. No será nuestra aspiración enumerarlas y describirlas. Nos limitaremos sólo al análisis de algunos accesos generales que pueden ser aplicados a la investigación de tales matrices.

Ya hace tiempo que en las matemáticas se acostumbra a comenzar el

estudio de las propiedades de las clases de objetos por la búsqueda de las invariantes y el análisis de los mismos. Una de las invariantes más importantes de cualquier matriz es el permanente de ésta. Al permanente le dedicamos un párrafo especial de este capítulo (4.4).

Estudiemos otras direcciones notables en este sentido. Ultimamente en las investigaciones teóricas de las matrices binarias se pasa, con una frecuencia cada vez mayor, del análisis de los tipos particulares de las matrices al estudio de sus clases. Se utilizan en dicho estudio las nociones de rango de frontera, de α -anchura y otras. Se denomina *traza* de una matriz a la suma de los elementos que constituyen su diagonal principal; el valor máximo para la traza, obtenido para toda clase de permutaciones de las filas y columnas de una matriz, recibe precisamente el nombre de *rango de frontera* (rango límite). Se llama α -anchura de la matriz a un número mínimo de columnas que pueden elegirse en ella de una manera tal que la suma de sus valores en cada fila sea no inferior a α .

He aquí uno de los procedimientos que se usa para elegir las clases de las matrices. Sea A una $(0, 1)$ -matriz de dimensión $m \times n$. Las sumas de sus elementos por filas la denotamos con r_i : $i = 1, 2, \dots, m$ y por columnas, con s_j : $j = 1, 2, \dots, n$. Es evidente que

$$\sum_{i=1}^m r_i = \sum_{j=1}^n s_j = \tau.$$

Haciendo uso de las totalidades de las sumas de las filas y de las sumas de las columnas de la matriz A , formamos los vectores correspondientes:

$$R = (r_1, r_2, \dots, r_m); \quad S = (s_1, s_2, \dots, s_n).$$

Al contrario, fijando los vectores R y S , cuyas componentes están representadas por los números no negativos, hallamos que ellos definen la clase de $(m \times n)$ -matrices A , la cual se denotará con $\mathcal{N}(R, S)$. Para las clases $\mathcal{N}(R, S)$ están determinadas las condiciones necesarias y suficientes del estado no vacío. Sin embargo, en esta parte del análisis combinatorio las investigaciones sólo acaban de iniciarse. Ni siquiera existe un método regular para calcular los números de todas las matrices de esta clase o, por lo menos, de las matrices normalizadas, es decir, tales en las que las líneas que constan solamente de ceros están tachadas y las líneas restantes están permutadas de un modo tal que

$$r_1 \geq r_2 \geq \dots \geq r_m > 0; \quad s_1 \geq s_2 \geq \dots \geq s_n > 0.$$

Para las clases de matrices se estudian los valores máximos y mínimos de los permanentes, de los rangos de frontera, de la α -anchura, etc.

4.2. RECTÁNGULOS Y CUADRADOS LATINOS

Así se denominan las disposiciones de los elementos por filas y por columnas, cuando los elementos en las líneas (tanto en las filas, como en columnas) no se repiten. La denominación "latinos" se debe, pro-

bablemente, a Euler quien usó, al estudiar la disposición de este tipo, las letras del alfabeto latino a título de sus elementos.

Sea dado un n -conjunto S . El problema de construcción sobre él (a base de sus elementos) de un $(r \times s)$ -rectángulo latino se reduce a la construcción de una tabla $(r \times s)$ -rectangular en la cual las líneas serán representadas, respectivamente, por las r - y s -permutaciones sin repetición de los elementos del conjunto S (evidentemente, $r, s \leq n$). Por ejemplo, la disposición

2	1	5	7	3
1	3	2	4	5
7	4	1	8	2
5	2	3	1	6

representará un (4×5) -rectángulo latino para $S = \{1, 2, \dots, 8\}$. En los $(r \times n)$ -rectángulos latinos las filas son permutaciones, elegidas de un modo tal que los elementos en las columnas no se repitan. Por ejemplo,

1	2	3	4	5	6
3	1	2	6	4	5
2	4	6	5	3	1
5	3	4	1	6	2

es un (4×6) -rectángulo latino para $S = \{1, 2, \dots, 6\}$. Si la primera fila de un rectángulo latino es tal que sus elementos quedan dispuestos en un orden fijado de antemano (con mayor frecuencia como $1, 2, \dots, n$), entonces el rectángulo se llama *normalizado*. El ejemplo que acabamos de aducir puede considerarse como ejemplo de rectángulo latino normalizado. Se analizan, a veces, rectángulos latinos que están normalizados no sólo por la primera fila, sino también por la primera columna.

Los rectángulos latinos que pueden obtenerse uno del otro mediante las permutaciones de las líneas y la reenumeración de los elementos se denominan *equivalentes*. Es fácil ver que para cada rectángulo latino existe un rectángulo equivalente normalizado por la primera fila y por la primera columna.

Resulta natural preguntar ¿cuántos rectángulos latinos del tipo dado pueden existir? Sea $L(r, n)$ el número de todos los $(r \times n)$ -rectángulos latinos y $N(r, n)$, el número de $(r \times n)$ -rectángulos latinos normalizados. Es obvio que

$$L(r, n) = n!N(r, n);$$

con la particularidad de que

$$L(1, n) = n!; \quad N(1, n) = 1.$$

Por cuanto los rectángulos latinos normalizados de dos filas son, en esencia, desórdenes, su número será igual al número de desórdenes:

$$N(2, n) = n! \left(\frac{1}{0!} - \frac{1}{1!} + \frac{1}{2!} - \dots + \frac{(-1)^n}{n!} \right) = D_n,$$

y, respectivamente,

$$L(2, n) = n! \quad N(2, n) = n! D_n.$$

Sin embargo, ya para el cálculo del número de rectángulos latinos de tres filas se necesitó superar dificultades mucho más grandes. Examinemos primeramente un tipo particular de tales rectángulos:

$$\begin{array}{ccccccc} 1 & 2 & 3 & 4 & \dots & n-1 & n \\ 2 & 3 & 4 & & \dots & n & 1 \\ a_1 & a_2 & a_3 & & \dots & & a_n \end{array}$$

y tratemos de determinar su número. El problema planteado de este modo es equivalente, en particular, al problema sobre el número de conmutadores, cuyos $2n$ jacks se disponen por un círculo, están unidos dos a dos y, además, ninguno de dos jacks vecinos están unidos entre sí. Este problema se conoce en otra interpretación, a saber, como problema sobre el número de colocaciones de n cónyuges en una mesa redonda bajo la condición de que ningún par conyugal se sienta juntos. La preocupación inicial es, naturalmente, sobre las damas; ellas se acomodan de $2n!$ maneras, quedándose vacantes las sillas a la derecha y a la izquierda de cada dama. Las sillas desocupadas se numeran $1, 2, \dots, n$. Ahora está claro que el primer conyuge no puede sentarse en las sillas primera y segunda; el segundo, en las sillas 2^a y 3^a , etc. hasta el n -ésimo cónyuge, para el cual están "prohibidas" las sillas 1^a y n -ésima. El problema se resuelve por la fórmula del método de inclusiones y exclusiones con las propiedades p_i : $a_i = i$, o bien $i + 1$ ($i = 1, 2, \dots, n-1$), $a_n = n$ ó a_1 . El número buscado es:

$$U_n = n! - 2n(n-1)! + \dots + (-1)^r \frac{2n}{2n-r} \binom{2n-r}{r} (n-r)! + \dots + (-1)^n \cdot 2.$$

En el caso general, el número de todos los rectángulos latinos de tres filas compuestos de n elementos se determina por la fórmula (véase [48]):

$$N(3, n) = \sum_{k=0}^{\left[\frac{n}{2}\right]} \binom{n}{k} D_{n-k} D_k U_{n-2k}.$$

Los números de $(3 \times n)$ -rectángulos latinos resultaron ser muy grandes. Los rectángulos latinos están por ahora débilmente estudiados, lo que se debe, probablemente, a la complejidad de las demostraciones y los enormes valores de sus números. En lo que se refiere a los rectángulos latinos con un número mayor de filas, podemos decir que en este dominio, en general, es poco lo que se conoce (véase [49]). En relación con esto no disminuye la importancia de las fórmulas asintóticas. Partiendo de los resultados conocidos:

$$L(2, n) = n! D_n \sim (n!)^2 e^{-1}; \quad L(3, n) \sim (n!)^3 e^{-3};$$

se ha supuesto que

$$L(k, n) \sim (n!)^k e^{-\binom{k}{2}}.$$

Esta hipótesis para $k < n^{1/3}$ fue demostrada por Yamamoto [50].

A pesar de que las tentativas de estimar superiormente el número de $(k \times n)$ -rectángulos latinos posibles no han llevado todavía a resultados exactos, ya podemos ver que este número es muy grande y crece rápidamente con el aumento de k . Esto se confirma con el hecho de que puede obtenerse la estimación inferior empleando la idea de ampliación de los rectángulos latinos.

Ampliar el rectángulo latino significa agregarle filas y columnas de tal modo que el rectángulo siga siendo latino. Esto sólo será posible, si el número de filas y columnas no sobrepasa el número de elementos n . La demostración de la posibilidad de ampliar un rectángulo se realiza por una simple repetición de los razonamientos que conducen a la adición de una fila por debajo y una columna, a la derecha. Como límite de ampliación interviene el cuadrado latino de orden n .

Sea dado un rectángulo latino $\mathcal{L}(r, n)$ de n elementos, los cuales se designarán con $1, 2, \dots, n$. Planteemos el problema de ampliar $\mathcal{L}(r, n)$ hasta que se obtenga $\mathcal{L}(r+1, n)$. Veamos las columnas $j = 1, 2, \dots, n$, y a cada j -ésima columna pongámosle en correspondencia un conjunto S_j compuesto de $n-r$ elementos que no han integrado la columna dada. Obtendremos una colección $M = (S_1, S_2, \dots, S_n)$, donde cada conjunto S_i consta de $n-r$ elementos.

El problema de ampliación de $\mathcal{L}(r, n)$ hasta $\mathcal{L}(r+1, n)$ se reduce ahora a la construcción de un s.r.d. para la colección M . Demostremos que tal s.r.d. existe. Efectivamente, tomemos $\mathcal{L}(r, n)$. Todo elemento i de n elementos se encuentra en él r veces. En los conjuntos $S_1, S_2, \dots, S_n$ el citado elemento se encuentra $(n-r)$ veces. Elijamos cualesquiera k conjuntos de $S_1, S_2, \dots, S_n$. En ellos habrá $k(n-r)$ elementos, contando también los que se repiten. Pero, cada elemento se encuentra no más de $n-r$ veces. Esto quiere decir que k conjuntos elegidos tienen no menos de k elementos distintos. La condición de existencia del s.r.d. está cumplida, lo que demuestra precisamente la posibilidad de la ampliación requerida.

La resolución del problema de ampliación de los rectángulos latinos permite dar una estimación inferior para el número de ellos. A saber, existen $n!$ rectangulares latinos $\mathcal{L}(1, n)$, cada uno de los cuales puede ser ampliado por lo menos hasta $(n-1)$ rectángulos latinos $\mathcal{L}(2, n)$. Por consiguiente, existen al menos $n!(n-1)!$ rectángulos $\mathcal{L}(2, n)$. Razonando de esta manera, llegaremos a la conclusión de que existe por lo menos

$$n!(n-1)! \dots (n-r+1)!$$

rectángulos latinos $\mathcal{L}(r, n)$.

Los cuadrados latinos en los que ahora concentraremos nuestra aten

ción representan un caso particular de los rectángulos latinos $\mathcal{L}(r, s)$ para $r = s = n$. En otras palabras, los cuadrados latinos son nada más que $(n \times n)$ -arreglos de n elementos, donde las filas y columnas son permutaciones de dichos elementos. Por ejemplo, todos los 4 cuadrados latinos posibles dos veces normalizados de cuarto orden son:

1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4
2 3 4 1	2 1 4 3	2 1 4 3	2 4 1 3
3 4 1 2	3 4 1 2	3 4 2 1	3 1 4 2
4 1 2 3	4 3 2 1	4 3 1 2	4 3 2 1

La existencia de los cuadrados latinos de cualquier orden se asegura no sólo por los razonamientos combinatorios generales, sino también por el hecho de que todo cuadrado latino puede interpretarse como la tabla de Cayley, es decir, la tabla de multiplicación en cierto grupo finito. Con el fin de calcular el número de cuadrados latinos y estimar este número superiormente hemos de superar, prácticamente, las mismas dificultades que tuvimos en el caso de los rectángulos latinos, en particular, las estimaciones inferiores tienen por expresión $(r - n)$:

$$n!(n - 1)! \dots 2!1!$$

Los cálculos del número de cuadrados latinos han mostrado que esta estimación es muy aproximada. Designemos con $L(n)$ el número de cuadrados latinos dos veces normalizados y con $I(n)$, el número de cuadrados latinos no equivalentes de orden n . Entonces, el número de todos los cuadrados latinos de n -ésimo orden es igual a $n!(n - 1)!L(n)$. Todos los números conocidos $I(n)$ y $L(n)$ se exponen en la siguiente tabla (véase [20]):

Tabla 4.1

n	$I(n)$				$L(n)$
1	1				1
2	1				1
3	1				1
4	4				2
5	56				2
6	9 408				22
7	16 942 080				563
8	535	281	401	856	1 676 257
9	377 597	570	964	258	816

En la teoría de los cuadrados latinos los problemas que mayor interés provocan son los relacionados con el concepto de ortogonalidad. Dos cuadrados latinos se llaman *ortogonales*, si al superponer uno de ellos sobre el otro se obtienen todos los pares de elementos ordenados sin repetición.

Dicho de otro modo, dos cuadrados latinos de orden n :

$$A_1 = [a_{ij}^{(1)}]; \quad A_2 = [a_{ij}^{(2)}]; \quad i, j = 1, 2, \dots, n,$$

se llaman ortogonales, si todos los n^2 pares $(a_{ij}^{(1)}, a_{ij}^{(2)})$ son distintos.

El estudio de la ortogonalidad de los cuadrados latinos tiene por origen el problema ampliamente conocido de Euler sobre 36 oficiales que llegaron de 6 regimientos. Cada regimiento asignaba 6 oficiales de rangos militares distintos. Se necesitaba incluir dichos oficiales en una formación de cuadro de un modo tal, que en ninguna línea hubiera repeticiones ni de rangos ni de nombres de regimientos. Dicho de otro modo, se planteaba el problema de construir dos cuadrados latinos ortogonales de sexto orden.

El problema sobre la construcción de un par de cuadrados latinos ortogonales de sexto orden no admitía resolución y Euler llegó a la conclusión de que tal par no existía. En una obra relacionada con este tema y publicada en 1782 Euler emitió la hipótesis de que la imposibilidad de hallar la solución se extendía también a los casos de $n = 10$, $n = 14$, y, en general, de $n \equiv 2 \pmod{4}$, es decir, $n = 2, 6, 10, 14, 18, \dots$.

Solamente pasados 118 años se dio la primera respuesta a la hipótesis de Euler. G. Tarry [51] confirmó la hipótesis a principios del siglo XX para $n = 6$, construyendo y arreglando una y otra vez todos los cuadrados latinos de sexto orden. La comprobación ulterior de la hipótesis se demoró de nuevo. Sólo para el año 1960 se demostró que los pares de cuadrados latinos ortogonales existían para cualquier orden, a excepción de $n = 2$ y $n = 6$, que resultaron ser la única excepción (véase [53]).

Como generalización natural del concepto de par de cuadrados latinos ortogonales interviene el conjunto compuesto por más de dos cuadrados latinos ortogonales dos a dos. Tal conjunto también se llama ortogonal. Reenumeremos los elementos de los cuadrados de una manera tal que quede normalizada la primera fila de cada cuadrado (al reenumerar los elementos de un cuadrado, la propiedad de ortogonalidad queda intacta). Entonces, los primeros elementos de las segundas filas deben ser diferentes y no iguales a 1. Por consiguiente, el número de cuadrados latinos ortogonales dos a dos de orden n no puede ser superior a $n - 1$. Un conjunto ortogonal se denomina *completo*, si en él están contenidos exactamente $n - 1$ cuadrados.

En la teoría de los cuadrados latinos ortogonales se distinguen dos grandes problemas: el problema de existencia de los pares y de las familias de cuadrados latinos ortogonales y el de su construcción. Estos dos problemas están estrechamente ligados, puesto que la demostración de la existencia con frecuencia también resulta ser construcción. El problema de existencia de los conjuntos de cuadrados latinos ortogonales resultó ser bastante difícil, y las demostraciones de los teoremas correspondientes, complejas y engorrosas. Damos a conocer aquí sólo algunos resultados.

1. No para todo cuadrado latino existe otro cuadrado latino que sea ortogonal al primero.

Este hecho, al igual que las condiciones de existencia, fue establecido en [52].

2. Si $n = p^\alpha$, donde p es un número primo y α , un número natural, entonces para $n \geq 3$ existe un conjunto completo de $n - 1$ cuadrados latinos ortogonales.

Demostración. Para cualquier $n = p^\alpha$ existe el campo de Galois. Supongamos que $a_0 = 0$ y $a_1 = 1, a_2, a_3, \dots, a_{n-1}$ son elementos del campo de Galois. Construyamos $n - 1$ matrices de orden n del modo siguiente:

$$A = [a_{ij}^{(l)}]; \quad i, j = 1, 2, \dots, n - 1; \quad l = 1, 2, \dots, n - 1$$

donde

$$a_{ij}^{(l)} = a_l a_i + a_j.$$

Esta familia de matrices A_l será precisamente el conjunto completo de cuadrados latinos ortogonales. En efecto:

a) cada una de las matrices construidas es un cuadrado latino, pues si cualquiera de ellas, digamos A_l , tiene dos elementos iguales en una misma fila, existen j y j' tales que

$$a_l a_i + a_j = a_l a_i + a_{j'}.$$

de donde

$$a_j = a_{j'}.$$

En cambio, si admitimos que dos elementos iguales figuran en una misma columna, entonces han de existir distintos i e i' tales que

$$a_l a_i + a_j = a_l a_{i'} + a_j;$$

y, por cuanto $a_l \neq 0$, tenemos $a_i = a_{i'}$, y, por lo tanto $i = i'$.

b) cada par de dichas matrices es ortogonal. Sea $1 \leq l < f \leq n - 1$, y supongamos que A_l no es ortogonal respecto de A_f , entonces existen i, i', j, j' tales que

$$(a_{ij}^{(l)}, a_{ij}^{(f)}) = (a_{i'j}^{(l)}, a_{i'j}^{(f)}),$$

es decir,

$$a_l a_i + a_j = a_l a_{i'} + a_{j'};$$

$$a_f a_i + a_j = a_f a_{i'} + a_{j'}.$$

De aquí

$$a_i(a_l - a_f) = a_{i'}(a_l - a_f),$$

y, por cuanto $a_l \neq a_f$, tenemos $a_i = a_{i'}$, de donde $i = i'$, y por lo tanto $j = j'$. El conjunto buscado está, de este modo, construido.

3. Enunciemos ahora una afirmación más general. Sean dados un número natural arbitrario n y su desarrollo en potencias naturales de los

distintos números primos: $n = p_1^{\alpha_1}, p_2^{\alpha_2}, \dots, p_N^{\alpha_N}$. Sea

$$t = \min(p_i^{\alpha_i} - 1); \quad t = 1, 2, \dots, N.$$

Si $t \geq 2$, existe una familia de t cuadrados latinos ortogonales de orden n .

Efectivamente, se ha demostrado anteriormente que para cada $n_i = p_i^{\alpha_i}$, $i = 1, 2, \dots, N$ existe una familia $n_i - 1$ (y por tanto, también de t) de cuadrados buscados. Para pasar al producto de las magnitudes se debe demostrar un teorema intermedio (la validez de la afirmación general se demuestra por la aplicación sucesiva de este teorema).

Teorema. Si se tienen dos familias, cada una de las cuales está compuesta por t cuadrados latinos ortogonales de órdenes n y n' respectivamente, existe una familia de t cuadrados latinos ortogonales de orden nn' .

Demostración. A toda familia de t cuadrados latinos ortogonales de orden n ($t \geq 2$, $n \geq 3$) se le puede poner en correspondencia una $[n^2 \times (t+2)]$ — tabla construida de los mismos n elementos (designados $1, 2, \dots, n$) de un modo tal que cualquier par de columnas contenga todas las n^2 2-muestras de elementos (tal tabla se llamará tabla de potencia 2). Permutemos en esta tabla sus filas de tal modo que en las primeras dos columnas (leyendo desde arriba) los pares se contengan en el siguiente orden: $(1, 1), (1, 2), \dots, (1, n), \dots, (n, 1), (n, 2), \dots, (n, n)$. En este caso, las siguientes t columnas reescritas por turno desde arriba (n columnas en una fila) forman t cuadrados latinos de orden n . De hecho, la primera columna de la tabla está construida de una manera tal que en ninguno de los t cuadrados latinos figuren en una fila dos elementos iguales. La segunda columna de la tabla está construida de tal manera que tampoco figuren dos elementos iguales en las columnas de los cuadrados. Por ejemplo, la (9×4) -tabla tiene la forma

1	1	1	1
1	2	2	2
1	3	3	3
2	1	2	3
2	2	3	1
2	3	1	2
3	1	3	2
3	2	1	3
3	3	2	1

corresponde a dos cuadrados latinos ortogonales:

1	2	3	1	2	3
2	3	1	3	1	2
3	1	2	2	3	1

Construyamos dos tablas correspondientes a dos t -familias de cuadrados latinos ortogonales de órdenes n y n' . Las dimensiones de estas tablas serán $n^2 \times (t+2)$ y $(n')^2 \times (t+2)$, respectivamente. Elijamos en estas tablas la i -ésima y j -ésima filas:

$$\begin{array}{cccc} a_{i,1} & a_{i,2} & \dots & a_{i,t+2}; \\ a'_{j,1} & a'_{j,2} & \dots & a'_{j,t+2} \end{array}$$

y construyamos, usando dichas filas, una fila compuesta por los pares de elementos:

$$a_{i,1}a'_{j,1}, a'_{i,2}a'_{j,2}, \dots, a_{i,t+2}a'_{j,t+2}.$$

De este modo, se obtienen en total $(nn')^2$ filas que forman una $[(nn')^2 \times (t+2)]$ -tabla en la que cada dos columnas forman todos los pares de los siguientes nn' elementos:

$$(1, 1), (1, 2), \dots, (1, n'), \dots, (n, 1), (n, 2), \dots, (n, n').$$

A esta tabla le corresponden t cuadrados latinos ortogonales de orden nn' .

La monografía [20] contiene una información abundante sobre las cuestiones examinadas en este párrafo.

4.3. BLOQUE-ESQUEMAS

El tipo de construcciones combinatorias del cual se trata en este párrafo, surgió en la literatura matemática aproximadamente a mediados del siglo pasado. J. Steiner publicó en 1853 un pequeño artículo [54] en el que se trataba sobre el problema siguiente: ¿qué número de símbolos pueden dividirse en ternas de un modo tal que cualquier par resulte ubicado en una y sólo una terna?

Se ha hallado sin dificultad alguna que sólo los números $n = 6k + 1$ y $n = 6k + 3$, es decir, $n \equiv 1, 3 \pmod{6}$ son la solución de este problema. En efecto, por cuanto cada elemento integra la terna junto con los otros dos, $n - 1$ debe ser un número par, y, por consiguiente, n , un número impar, es decir, $n \equiv 1 \pmod{2}$. Además, en cada terna hay tres pares. Cada par aparece una sola vez. En total hay $\frac{n(n-1)}{2}$ pares. Este número ha de ser múltiplo de tres. De aquí, $n \equiv 1, 3 \pmod{6}$, es decir, $n = 3, 7, 9, 13, 15, 19, \dots$. Estas condiciones son necesarias. Steiner planteó una cuestión: ¿son ellas también suficientes? Reiss [55] dio en 1859 una respuesta afirmativa.

Ni Steiner ni Reiss sabían que un problema combinatorio análogo fue planteado y resuelto un poco antes, en 1847, por Kirkman [56]. Pasados tres años, Kirkman propuso otro problema combinatorio [57]: una profesora saca cada día a pasear a un grupo de 15 niñas; las forma en 5 filas de a 3 niñas en cada una; se requiere organizar las ternas de un modo tal que en el transcurso de 7 días cada alumna se encuentre con todas las demás una sola vez.

El sistema de ternas descrito en este problema lleva el nombre de Kirkman. En comparación con el sistema de ternas de Steiner, en éste está presente una condición adicional (la llamada condición de solubilidad): el conjunto de ternas ha de ser partido en $\frac{n-1}{2}$ componentes (en el problema de Kirkman hay 7 componentes y cada una corresponde a un día

de la semana) de tal modo que cada uno de n elementos aparece exactamente una vez en las ternas de cada componente.

Surge la pregunta: ¿cuántos sistemas del tipo dado existen? Para las ternas de Steiner se sabe lo siguiente. Si $n = 3, 7, 9$, entonces existe solamente un tipo de tal sistema, a saber:

a) para $n = 3$, solamente el sistema de una terna

1 2 3

b) para $n = 7$, solamente de siete ternas

1	2	3								
1	4	5		2	4	6		3	4	7
1	6	7		2	5	7		3	5	6

c) para $n = 9$, solamente el sistema de 12 ternas

1	2	3								
1	4	5		2	4	9		3	4	8
1	6	8		2	5	6		3	5	7
1	7	9		2	7	8		3	6	9
									4	6
									5	8
									7	9

Resultó que para $n = 13$ existen dos sistemas de soluciones no isomorfas, es decir tales que no pueden ser obtenidos uno del otro por sustitución de los elementos o por permutación de las ternas. Estos sistemas constan de 22 ternas iguales

1	2	3												
1	4	5		2	4	6								
1	6	7		2	5	7		4	3	8		7	3	11
1	8	9		2	8	10		4	7	9		7	8	13
1	10	11		2	9	12		4	10	13		7	10	12
1	12	13		2	11	13		4	11	12				
													8	5
													8	6
													6	9
													3	5

y cuatro no equivalentes: para el primer sistema

3	6	10		5	6	13
3	9	13		5	9	10

y para el segundo,

3	6	13		5	6	10
3	9	10		5	9	13

Para el año 1925 fueron terminados los cálculos de los sistemas de ternas de Steiner para $n = 15$; tales sistemas diferentes resultaron ser 80. Por lo que sabe, no se volvieron a calcular las ternas de Steiner para $n > 15$. En lo que atañe al número de sistemas no isomorfos de ternas de Steiner, éste crece indefinidamente con el aumento de n (véase [58]).

La solución del problema de Kirkman (conocida desde el año 1921) es

1	2	5	1	3	9	1	4	15	1	6	11
3	14	15	2	8	15	2	9	11	2	7	12
4	6	12	4	11	13	3	10	12	3	8	13
7	8	11	5	12	14	5	7	13	4	9	14
9	10	13	6	7	10	6	8	14	5	10	15
1	8	10	1	7	14	1	12	13			
2	13	14	2	4	10	2	3	6			
3	4	7	3	5	11	4	5	8			
5	6	9	6	13	15	7	9	15			
11	12	15	8	9	12	10	11	14			

Los métodos de construcción de los sistemas de ternas son, en lo principal, recursivos, es decir, se construyen sistemas "iniciales" de ternas para los valores pequeños de los parámetros, y después se emplean los métodos que se apoyan en esta reserva que ya se tiene. He aquí algunos de estos métodos.

Si existen dos sistemas de ternas de Steiner, A y B , de órdenes v_1 y v_2 , respectivamente, puede construirse un sistema de Steiner más, el C , de orden $v = v_1 \cdot v_2$, que contiene subsistemas isomorfos a los dados.

Sea (a_i, a_j, a_k) una terna del sistema A y sea (b_r, b_s, b_u) una terna del sistema B . Construyamos el sistema de ternas C a base de los elementos c_{ij} ($i = 1, 2, \dots, v_1; j = 1, 2, \dots, v_2$) que forman las ternas (c_{ir}, c_{js}, c_{ku}) , si está cumplida alguna de las condiciones siguientes: 1) $r = s = u$, $(a_i, a_j, a_k) \in A$; 2) $i = j = k$, $(b_r, b_s, b_u) \in B$; 3) $(a_i, a_j, a_k) \in A$, $(b_r, b_s, b_u) \in B$. Un sistema isomorfo al sistema A será el primero de ellos a condición de que: $r = s = u = 1$, y de sistema isomorfo a B servirá el segundo con la condición de que $i = j = k = 1$.

En 1893 Moore [59] propuso el método de construcción de las ternas de Steiner basado en las ideas expuestas más arriba.

Sean dados los sistemas de Steiner: S_2 de orden v_2 , que está integrado por S_3 de orden v_3 , y S_1 de orden $v_1 > 1$; entonces, se puede construir un sistema de Steiner S de orden $v = v_3 + v_1(v_2 - v_3)$ que tenga subsistemas isomorfos a S_1, S_2, S_3 .

Tomemos un conjunto de $v = v_3 + v_1(v_2 - v_3)$ elementos y representémoslo en la forma siguiente

$$\begin{array}{ll}
 M_0 : a_1 & a_2 \dots a_{v_3}; \\
 M_1 : b_{11} & b_{12} \dots b_{1, v_2 - v_3}; \\
 M_2 : b_{21} & b_{22} \dots b_{2, v_2 - v_3}; \\
 \dots\dots\dots \\
 M_{v_1} : b_{v_1, 1} b_{v_1, 2} \dots b_{v_1, v_2 - v_3}.
 \end{array}$$

Formemos las ternas rigiéndonos por las siguientes reglas.

1. Pongamos los elementos de M_0 en una correspondencia biunívoca con los elementos de S_3 y elijamos todas las ternas (a_i, a_j, a_k) , si ellas corresponden a las ternas del sistema S_3 .

2. Pongamos los elementos del par de conjuntos M_0 y M_i , $i = 1, 2, \dots, v_1$ en una correspondencia biunívoca con el sistema S_2 de un modo tal que M_0 corresponda a su subsistema S_3 . Las ternas compuestas por los elementos de M_0 ya se han encontrado según la regla antecedente. Las demás ternas han de contener no más de un elemento de M_0 , y, por consiguiente, tienen la forma: (a_m, b_{ij}, b_{ik}) ó (b_{ij}, b_{ik}, b_{ir}) . Por supuesto, ellas deben corresponder a las ternas restantes del sistema S_2 .

3. Definamos el sistema S_1 sobre los números $1, 2, \dots, v_1$. Si (j, i, r) es una terna de este sistema, formemos todas las ternas de la forma (b_{jx}, b_{ky}, b_{rz}) , donde $x + y + z \equiv 0 \pmod{v_2 - v_3}$.

Las reglas citadas dan todas las ternas del sistema S de orden $v = v_3 + v_1(v_2 - v_3)$.

En lo que se refiere a las ternas de Kirkman, podemos decir que se han encontrado métodos, con ayuda de los cuales ellas se construyen para los órdenes (véase [60])

$$v = 15, 15 \cdot 3^n, 3^n, 2^{2n} - 1.$$

Pueden considerarse también ternas de otros tipos, diferentes de las de Steiner y Kirkman. Así, por ejemplo, un cuadrado latino genera el conjunto de n^2 ternas ordenadas, donde cada terna (i, j, k) corresponde a que en la i -ésima fila y la j -ésima columna se dispone un elemento k ; $i, j, k = 1, 2, \dots, n$. La condición de que una totalidad de ternas (i, j, k) ($i, j, k = 1, 2, \dots, n$) forma un cuadrado latino consiste en que cualquier par de componentes de la terna ha de encontrarse una y sólo una vez.

Introducidos los sistemas de ternas de Steiner y Kirkman, los problemas en el dominio de las muestras combinatorias, a lo que parece, deberían concentrarse en las siguientes direcciones:

a) en la elaboración de los métodos de construcción eficaz de los diversos sistemas, en primer lugar, de las ternas de Steiner y Kirkman;

b) en las generalizaciones de los problemas de ternas para los sistemas de 4-, 5-, 6-, etc. conjuntos. Aunque dichos problemas siguen siendo actuales hoy día (véase [61]), tal camino de desarrollo paulatino fue perturbado. Tuvo lugar un salto lógico, cosa no tan rara en las matemáticas. La superación de las dificultades se empezó a buscar en la vía de construcción de una teoría más general.

Actualmente entre los sistemas de incidencia los que mayor interés presentan son los bloque-esquemas (esquemas en bloque). El estudio de estos se efectúa con gran actividad, lo que se explica tanto por las necesidades de la práctica, como por los intereses científicos puros.

En forma simplificada el origen de los bloque-esquemas puede ser descrito del modo siguiente. Los resultados del experimento realizado a un mismo tiempo en condiciones determinadas se escriben en forma de columna (bloque). Los resultados reiterativos, obtenidos en otras condiciones (cambiadas), proporcionan bloques nuevos. La colección de los bloques da

una tabla bidimensional llamada bloque-esquema (en forma abreviada BD, que proviene de la expresión inglesa "block-design").

De la elaboración estática de los resultados nace un enfoque al problema de planificación del experimento, es decir, de formación de su esquema (plan) con el fin de obtener resultados óptimos con un número mínimo de pruebas.

Para el análisis combinatorio, entre un gran número de problemas concernientes a los bloque-esquemas los que mayor interés representan son los problemas de existencia y clasificación, los que estudian las propiedades de los diferentes tipos de bloque-esquemas, los métodos de construcción de los mismos, la correlación, la elaboración de las cuestiones contiguas de la matemática discreta, de las geometrías finitas, de los grupos finitos, etc. Estos problemas teóricos se analizan intensamente desde el fin de los años 30 y el principio de los 40.

Sea un v -conjunto de elementos $m_1, m_2, \dots, m_v$. Los elementos de dicho conjunto están distribuidos entre b bloque-subconjuntos $M_1, M_2, \dots, M_b$, la intersección de los cuales no es obligatoriamente un vacío. El número de elementos en el bloque M_j lleva el nombre de volumen del bloque y se denota con k_j . Los elementos pueden aparecer en varios bloques. Sea r_i el número de apariciones del elemento m_i (es decir, el número de bloques que contienen m_i), $i = 1, 2, \dots, v$; por fin, introduzcamos el número de repeticiones (pares no ordenados) de los elementos: λ_p ($p = 1, 2, \dots, (v/2)$). En tal caso suele decirse que el conjunto de bloques $M_1, \dots, M_b$ forma un *bloque-esquema* con los parámetros $v, b, r_i, k_j, \lambda_p$ (si en lugar de los pares se analizan las ternas o l -muestras de elementos, entonces la disposición correspondiente se denomina *configuración táctica*). Así pues, un bloque-esquema se caracteriza por los parámetros $v, b, r_i, k_i, \lambda_p$.

Pueden haber varios tipos diferentes de bloque-esquemas. Por eso introduzcamos el principio de la clasificación de los bloque-esquemas tal como se formó hasta el presente. Ante todo, un bloque-esquema se llama *completo*, si en cada bloque están contenidos todos los elementos del conjunto, es decir, $k_j = v, j = 1, 2, \dots, b$, e *incompleto*, si $k_j < v, j = 1, 2, \dots, b$. Los cuadrados latinos pueden servir de ejemplo de los bloque-esquemas completos.

Los bloque-esquemas incompletos se dividen, a su vez, en dos grandes clases: la clase de bloque-esquemas incompletos equilibrados (en la forma abreviada BIB-esquemas, lo que viene del inglés "balanced incomplete block design") y la clase de bloque-esquemas incompletos parcialmente equilibrados (PBIB, del inglés "partially balanced incomplete block design").

Un bloque-esquema se llama BIB-esquema, si $k_1 = k_2 = \dots = k_b = k < v; r_1 = \dots = r_v = r$, y $\lambda_p = \lambda$ para cualquier p ; dicho de otro modo, en un BIB-esquema todos los elementos tienen un mismo número de repeticiones, todos los bloques son de igual volumen y cada

par de elementos aparece en un mismo número $\lambda = \text{const}$ de bloques. Los PBIB-esquemas se diferencian en que en ellos $\lambda \neq \text{const}$. Debido a esta circunstancia se genera un gran número de diferentes tipos de PBIB-esquemas, lo que crea dificultades considerables en la investigación de ellos.

Se denomina bloque-esquema incompleto parcialmente equilibrado con m clases, o, en forma más breve, PBIB(m)-esquema a un bloque-esquema, en el que el conjunto de todos los $\binom{v}{2}$ pares de sus elementos puede ser partido en m clases disjuntas (un par de elementos referente a la i -ésima clase se denominará i -conexo), de un modo tal que

a) cualquier par de elementos i -conexos esté contenido exactamente en λ_i bloques ($i = 1, 2, \dots, m$);

b) para cualquier elemento exista exactamente n_i elementos que sean i -conexos con el primero ($i = 1, 2, \dots, m$);

c) para todo par de elementos i -conexos, λ y β , el número de elementos, j -conexos con α , y, al mismo tiempo, l -conexos con β , es igual a p_{jl}^i , con la particularidad de que $p_{jl}^i = p_{lj}^i$ ($i, j, l = 1, 2, \dots, m$).

Por consiguiente, los parámetros de un PBIB(m)-esquema son los números

$$v, b, k, r, \lambda_i, n_i, p_{jl}^i \quad (i, j, l = 1, 2, \dots, m).$$

Con el objeto de contribuir a concebir un gran número de construcciones combinatorias nuevas y, a veces, complejas, proponemos al lector el esquema de la clasificación primaria de los bloque-esquemas señalando las continuaciones posibles (fig. 4.1).

La teoría moderna de los bloque-esquemas es enorme en cuanto al número y volumen de las investigaciones reales, a menudo dispersadas y débilmente organizadas. Más abajo se expondrán los fundamentos de esta teoría sin pretender interpretar los hechos plenamente.

Al principio de este capítulo ya se ha considerado en forma general el método de representación matricial de la estructura de un conjunto. Apliquémoslo a los bloque-esquemas. Sea un bloque-esquema con los parámetros $v, b, k, r^{1)}$. Pongamos

$$n_{ij} = \begin{cases} 1, & \text{si el } i\text{-ésimo elemento se encuentra en el } j\text{-ésimo bloque} \\ 0, & \text{en el caso contrario,} \end{cases}$$

$$i = 1, 2, \dots, v; \quad j = 1, 2, \dots, b.$$

La $(v \times b)$ -matriz $N = \|n_{ij}\|$ se llama matriz de incidencia del bloque-esquema dado; en este caso cada fila de la matriz N contiene r unidades y cada columna contiene k unidades (λ_p es igual al producto escalar de

¹⁾ No vamos a escribir el parámetro λ , si no se indica de qué, precisamente, bloque-esquema (equilibrado o parcialmente equilibrado) se trata en el caso a considerar.

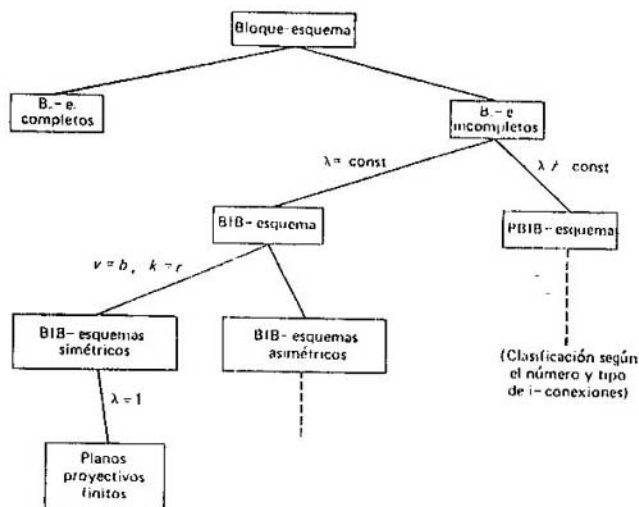


Fig.4.1.

dos filas). Por ejemplo, un BIB-esquema con los parámetros $v = 6$, $b = 10$, $k = 3$, $r = 5$, $\lambda = 2$ tiene la siguiente representación matricial (matriz de incidencia):

Bloque-esquema

1	1	1	1	1	2	3	2	3	2
2	3	4	5	2	3	4	4	5	4
3	4	5	6	6	5	6	5	6	6

Matriz

1	1	1	1	1	0	0	0	0	0
1	0	0	0	1	1	0	1	0	1
1	1	0	0	0	1	1	0	1	0
0	1	1	0	0	0	1	1	0	1
0	0	1	1	0	1	0	1	1	0
0	0	0	1	1	0	1	0	1	1

He aquí un ejemplo más: supongamos que para un 9-conjunto se examina el siguiente sistema de subconjuntos:

$$\begin{aligned}
 B_1 &= (1, 2, 3); & B_2 &= (4, 5, 6); & B_3 &= (7, 8, 9); & B_4 &= (1, 4, 7); \\
 B_5 &= (2, 5, 8); & B_6 &= (3, 6, 9); & B_7 &= (1, 5, 9); & B_8 &= (2, 6, 7); \\
 B_9 &= (3, 4, 8); & B_{10} &= (1, 6, 8); & B_{11} &= (2, 4, 9); & B_{12} &= (3, 5, 7).
 \end{aligned}$$

No es difícil ver que este sistema es un BIB-esquema con los parámetros $v = 9$, $b = 12$, $k = 3$, $r = 4$, $\lambda = 1$. Escribamos el bloque-esquema, como antes, en forma de una $(k \times b)$ -tabla rectangular:

1	4	7	1	2	3	1	2	3	1	2	3
2	5	8	4	5	6	5	6	4	6	4	5
3	6	9	7	8	9	9	7	8	8	9	7

La matriz de incidencia de este bloque-esquema es

1	0	0	1	0	0	1	0	0	1	0	0
1	0	0	0	1	0	0	1	0	0	1	0
1	0	0	0	0	1	0	0	1	0	0	1
0	1	0	1	0	0	0	0	1	0	1	0
0	1	0	0	1	0	1	0	0	0	0	1
0	1	0	0	0	1	0	1	0	1	0	0
0	0	1	1	0	0	0	1	0	0	0	1
0	0	1	0	1	0	0	0	1	1	0	0
0	0	1	0	0	1	1	0	0	0	1	0

Pese a que la matriz parece ser engorrosa, las matrices binarias de incidencia son en la práctica muy simples para las operaciones. Del análisis de sus estructuras, podemos obtener directamente una serie de propiedades, incluidas las relaciones entre los parámetros. Por ejemplo, para los bloque-esquemas tenemos en general

$$\sum_{i=1}^n r_i = \sum_{j=1}^h k_j.$$

lo que es obvio. En particular, para los BIB-esquemas tenemos

$$bk = vr; \quad (1)$$

$$r(k-1) = \lambda(v-1). \quad (2)$$

La primera de las relaciones expresa un simple cálculo del número de elementos en el bloque-esquema (es decir, el número de todas las unidades en la matriz de incidencia). En la segunda relación los razonamientos son los siguientes: cada elemento se encuentra r veces y está siempre contiguo de $k-1$ elementos. Este mismo elemento se encuentra λ veces haciendo un par con cada uno de las $v-1$ elementos.

Veamos las relaciones entre los parámetros del PBIB(m)-esquema. Es evidente que subsiste, como antes, la igualdad (1). Establezcamos algunas otras igualdades:

$$\sum_{i=1}^m n_i = v-1; \quad \sum_{i=1}^m \lambda_i n_i = r(k-1);$$

$$\sum_{i=1}^m p_{ji}^i = \begin{cases} n_j, & \text{si } i \neq j; \\ n_j - 1, & \text{si } i = j; i, j = 1, 2, \dots, m; \end{cases} \quad (3)$$

$$n_i p_{ji}^j = n_j p_{ji}^i = n_i p_{ji}^j; \quad i, j, l = 1, 2, \dots, m.$$

La primera de las relaciones es obvia. La segunda es una modificación de (2) y se demuestra por razonamientos análogos. En la siguiente igualdad

$\sum_{i=1}^m p_{ji}^i$ significa el número de elementos j -conexos con cualquiera de los dos elementos i -conexos. Es igual, por consiguiente, a n_j (si $i \neq j$), o bien a $n_j - 1$, (si $i = j$). En lo que se refiere a la última relación, esta se

demuestra del modo siguiente: tomemos un elemento fijo α . Por definición, existen n_i elementos $\beta_1, \beta_2, \dots, \beta_{n_i}$, i -conexos con α , y, además, n_j elementos $\gamma_1, \gamma_2, \dots, \gamma_{n_j}$, j -conexos con α . Construyamos una tabla binaria

$$\begin{pmatrix} b_{\gamma_1, \beta_1}^{\beta_1} & \dots & b_{\gamma_1, \beta_{n_i}}^{\beta_{n_i}} \\ \dots & \dots & \dots \\ b_{\gamma_{n_j}, \beta_1}^{\beta_1} & \dots & b_{\gamma_{n_j}, \beta_{n_i}}^{\beta_{n_i}} \end{pmatrix}$$

en la cual

$$b_{\gamma, \beta}^{\beta} = \begin{cases} 1, & \text{si } \beta_i \text{ y } \gamma_j \text{ son } i\text{-conexos;} \\ 0, & \text{en el caso contrario.} \end{cases}$$

Calculemos el número de unidades en cada columna de la matriz:

$$\sum_{\gamma=1}^{n_j} b_{\gamma, \beta}^{\beta} = p_{j, i}^i;$$

y en cada columna:

$$\sum_{\beta=1}^{n_i} b_{\gamma, \beta}^{\beta} = p_{i, j}^j;$$

De los dos métodos de cálculo de todas las unidades en la matriz obtenemos

$$n_i p_{j, i}^i = n_j p_{i, j}^j.$$

Las relaciones (1), (2) y (1), (3) son condiciones necesarias para que exista los BIB- y PBIB(m)-esquemas, respectivamente.

Sea $N = [n_{ij}]$ una $(v \times b)$ matriz de incidencia de un bloque-esquema incompleto equilibrado con los parámetros v, b, r, k, λ . En este caso N satisface la siguiente ecuación matricial

$$NN^T = (r - \lambda)I + \lambda J,$$

donde N^T es la matriz transpuesta de N ; I es la matriz unidad de orden v ; J es una matriz integrada plenamente por unidades, también de orden v .

Efectivamente, calculemos el producto

$$N \cdot N^T = B = [b_{ij}]; \quad i, j = 1, 2, \dots, v.$$

El elemento b_{ij} es igual al producto escalar de la i -ésima fila de N por la j -ésima columna de N^T , por consiguiente,

$$b_{ij} = \sum_{l=1}^b n_{il} n_{jl} = \begin{cases} r, & \text{si } i = j \\ \lambda, & \text{si } i \neq j; \end{cases}$$

$$i, j = 1, 2, \dots, v.$$

De aquí

$$B = \begin{pmatrix} r & & \\ & r & \lambda \\ & \lambda & \dots r \end{pmatrix} = rJ + \lambda(J - I) = (r - \lambda)I + \lambda J.$$

Para la matriz B puede hallarse el valor de su determinante:

$$\det B = (r - \lambda)^{v-1} (v\lambda - \lambda + r).$$

Para obtener este resultado basta sustraer la primera columna de todas las demás columnas, y a continuación, agregar la suma de todas las filas, partiendo de la segunda, hacia la primera fila. En este caso, por arriba de la diagonal principal se dispondrán sólo los ceros: $a'_{11} = r + (v - 1)\lambda$; $a'_{22} = \dots = a'_{vv} = r - \lambda$, lo que nos da el resultado buscado. Por cuanto de que $k < v$ se deduce que $\lambda = r \frac{k-1}{v-1} < r$, entonces $\det B \neq 0$. Haciendo uso de (2), tenemos

$$\det B = (r - \lambda)^{v-1} rk \neq 0.$$

De este modo, el rango de la matriz B es igual a v . Por otra parte, el rango de N (y el de N^T) no sobrepasa b . Según se sabe por la teoría de las matrices, el rango de un producto no es superior a los rangos de los factores. De aquí se deduce una desigualdad importante

$$b \geq v,$$

llamada desigualdad de Fisher, la cual sirve también de condición necesaria para la existencia de un BIB-esquema.

Del análisis de la matriz de incidencia N y de la matriz $B = N \cdot N^T$ pueden deducirse una serie de teoremas referentes a los tipos particulares de los bloque-esquemas (véase [62]).

Las matrices de incidencia no son el único método de describir los bloque-esquemas; otro método análogo es el de aplicación de las formas cuadráticas. Este último método es singularmente interesante, porque la teoría de las formas cuadráticas está suficientemente elaborada. En efecto, la correlación $N \cdot N^T = (r - \lambda)I + \lambda J$ se reencuentra en las formas cuadráticas. Introduzcamos las variables indeterminadas: $x_1, x_2, \dots, x_v$, correspondientes a los elementos de los bloque-esquemas, y asignamos a cada bloque b_j la forma

$$L_j = \sum_{i=1}^v n_{ij} x_i, \quad j = 1, 2, \dots, v,$$

donde n_{ij} es la multiplicidad del elemento i en el bloque b_j . Entonces, a la ecuación matricial $N \cdot N^T = (r - \lambda)I + \lambda J$ se le pondrá en correspondencia la forma cuadrática:

$$\sum_{i=1}^b L_i^2 = (r - \lambda) \sum_{i=1}^v x_i^2 + \lambda \left(\sum_{i=1}^v x_i \right)^2.$$

En la teoría del análisis combinatorio y en las aplicaciones del mismo los bloque-esquemas desempeñan un papel de importancia. Varias de las estructuras combinatorias conocidas representan tipos particulares de los

bloque-esquemas o son equivalentes a los últimos. Aclaremos esto con un ejemplo.

Según lo dicho con anterioridad, los cuadrados latinos resultan ser bloque-esquemas completos. Las ternas de Steiner son, en esencia, BIB-esquemas, para los cuales $k = 3$, $\lambda = 1$. Para ellos las correlaciones (1) y (2) entre parámetros adquieren la forma

$$3b = vr; \quad 2r = v - 1;$$

de donde

$$v = 2r + 1; \quad b = \frac{r(2r + 1)}{3}$$

y, por consiguiente,

$$r \equiv 0, 1 \pmod{3}; \quad v \equiv 1, 3 \pmod{6}.$$

Las matrices de Hadamard H son equivalentes a los tipos particulares de los bloque-esquemas. Efectivamente, tomemos la matriz H de orden $n \geq 3$. Se conoce (véase § 4.1) que $n \equiv 0 \pmod{4}$. Normalicemos H , es decir, multiplicando las filas y las columnas correspondientes por -1 , reduzcámosla a una forma, en la cual la primera fila y la primera columna estarán compuestas de unidades positivas. Tal matriz H de orden $n = 4t \geq 8$ resulta ser equivalente al BIB-esquema simétrico con los parámetros

$$v = b = 4t - 1; \quad k = r = 2t - 1; \quad \lambda = t - 1.$$

En efecto, a toda matriz normalizada H se le puede poner en correspondencia una $(0, 1)$ -matriz A de orden $v = 4t - 1$, en cada línea de la cual se tienen $k = 2t - 1$ unidades. El método de esta operación consiste en que en la matriz normalizada H se tachan la primera fila y la primera columna, y en la matriz restante todos los menos uno se sustituyen por ceros. La matriz A satisface la relación

$$A \cdot A^T = tI + (t - 1)J$$

y representa, pues, una matriz de incidencia de un bloque-esquema incompleto equilibrado en el que $r - \lambda = t$, y $\lambda = t - 1$. Los razonamientos son fácilmente invertidos, lo que precisará demostrará nuestra afirmación.

En el § 4.2, al demostrar el teorema de existencia de la familia de cuadrados latinos ortogonales, se ha mostrado que la construcción de tal familia es equivalente a la conformación de una tabla ortogonal de fuerzas. Resulta, no obstante, que la existencia de la última origina la existencia de un bloque-esquema de determinado tipo. Expliquemos este hecho más detalladamente, pero, antes demos a conocer la definición general de la tabla ortogonal.

Sea dada una $(m \times N)$ -matriz A con elementos del conjunto $S = \{0,$

1, 2, ..., $s - 1$). Existen en total s^t diferentes columnas de longitud t con los elementos del conjunto mencionado. Si en cada submatriz de t filas de la matriz A la columna $(x_1, x_2, \dots, x_t)^T$ aparece $\lambda(x_1, x_2, \dots, x_t)$ veces, donde el número λ es positivo y no varía al realizarse las permutaciones de $x_1, x_2, \dots, x_t$, entonces la matriz A se llama *tabla parcialmente equilibrada de fuerza t* con N columnas, m filas, s símbolos y los parámetros $\lambda(x_1, x_2, \dots, x_t)$, o, en una forma más breve, tabla (N, m, s, t) con los parámetros $\lambda(x_1, x_2, \dots, x_t)$. Si $\lambda = \text{const}$ para todas las columnas $(x_1, x_2, \dots, x_t)$, entonces la tabla (N, m, s, t) se denomina *tabla ortogonal de índice λ* .

Aduzcamos, a título de ejemplo, una tabla ortogonal $(18, 7, 3, 2)$ de índice $\lambda = 2$ con los símbolos 0, 1, 2:

Tabla 4.2

Números de las filas	Números de las columnas																	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	0	1	2	0	1	2	0	1	2	0	1	2	0	1	2	0	1	2
2	0	1	2	0	1	2	1	2	0	2	0	1	1	2	0	2	0	1
3	0	1	2	1	2	0	0	1	2	2	0	1	2	0	1	1	2	0
4	0	1	2	2	0	1	2	0	1	0	1	2	1	2	0	1	2	0
5	0	1	2	1	2	0	2	0	1	1	2	0	0	1	2	2	0	1
6	0	1	2	2	0	1	1	2	0	1	2	0	2	0	1	0	1	2
7	0	0	0	0	0	0	1	1	1	1	1	1	2	2	2	2	2	2

Si separamos de esta tabla las tres primeras columnas y la última fila obtendremos una tabla parcialmente equilibrada $(15, 6, 3, 2)$ con los parámetros

$$\lambda(x_1, x_2) = \begin{cases} 2, & \text{si } x_1 \neq x_2; \\ 1, & \text{si } x_1 = x_2. \end{cases}$$

La existencia de k cuadrados latinos biunívocamente ortogonales de orden s es equivalente a la existencia de una tabla ortogonal $(s^2, k + 2, s, 2)$ de índice $\lambda = 1$.

Efectivamente, sea que tenemos una tabla ortogonal $(s^2, k + 2, s, 2)$ y supongamos que los elementos en esta tabla están designados con los números 1, 2, ..., s . Permutemos las columnas en la tabla de una manera tal que los elementos de las columnas en las primeras dos filas se encuentren en orden

$$(1, 1), (1, 2), \dots, (1, s), \dots, (s, 1), (s, 2), \dots, (s, s).$$

Para cada $l = 3, 4, \dots, k + 2$ construyamos la tabla $A_l = (s^2, s, s, 2)$ del

modo siguiente: la primera columna A_i consta de los primeros s elementos de la fila i de la tabla inicial; la segunda columna, de los siguientes s elementos, etc. hasta la última compuesta por los últimos s elementos de la misma fila. Las tablas $A_1, A_2, \dots, A_{k+2}$, conformadas del modo indicado, forman un conjunto compuesto por k cuadrados latinos biunívocamente ortogonales de orden s . El hecho de que la tabla A_i no contiene dos elementos iguales en la columna lo muestra el examen de la primera fila de la tabla inicial. Una propiedad análoga de las filas se ve del examen de la segunda fila. Si $i \neq j$, entonces, A_i y A_j son ortogonales, lo que se desprende del método de construcción de las filas correspondientes. La demostración como es fácil comprobar es convertible.

Las investigaciones científicas en el dominio de los bloque-esquemas se agrupan actualmente alrededor del problema de existencia (o no existencia) de los tipos aislados de los bloque-esquemas, del estudio de sus propiedades y de la búsqueda de los métodos de su construcción. La teoría combinatoria se encuentra en este respecto sólo al principio de su desarrollo. Para el estudio ulterior de este tema podemos recomendar al lector las capítulos 10—16 del libro [4].

Las investigaciones modernas en el dominio de las disposiciones combinatorias y de los bloque-esquemas requieren, en particular, la utilización de la teoría de los números, la teoría de las matrices, la teoría de los cuerpos convexos, etc.

4.4. PERMANENTES

Este concepto ha adquirido en el análisis combinatorio moderno una importancia teórica y práctica tan grande que consideramos necesario dedicarle un párrafo especial aparte.

Sean dados un conjunto $S = \{s_1, \dots, s_n\}$ y cierta familia de sus subconjuntos $M = \{S_1, S_2, \dots, S_m\}$. El teorema de P. Hall da las condiciones necesarias y suficientes para que la familia M posea sistemas de representantes distintos y una frontera inferior para el número de tales sistemas. Sin embargo, en muchos problemas del análisis combinatorio surge la necesidad de obtener estimaciones más exactas de este número. Este objetivo se consigue por introducción de una nueva noción, permanente de la matriz, y estudiando las propiedades del mismo.

Sea una matriz $A = \|a_{ij}\|$, $i = 1, 2, \dots, m$; $j = 1, 2, \dots, n$, cuyos elementos están representados por números reales. Se denomina *permanente* de la matriz citada la expresión (el número)

$$\text{per } A = \sum_{\sigma} \prod_{i=1}^m a_{i\sigma(i)},$$

en la que la sumación se realiza respecto de todos los encajes σ de $\{1, 2, \dots, m\}$ en $\{1, 2, \dots, n\}$. En este caso se supone, naturalmente, que $m \leq n$.

Sea A una matriz de incidencia para M respecto de S , entonces per A será el número de sistemas de representantes distintos, de los cuales dispone la familia de subconjuntos M .

Los permanentes fueron introducidos en 1812, casi simultáneamente, en las memorias de Binet [63] y Cauchy [64] en relación con el desarrollo de la teoría de determinantes como tipo especial de las funciones simétricas de signo variable. El término "permanente" apareció por primera vez en la obra de Muir [65], donde están enunciadas algunas propiedades del permanente.

Es evidente que el permanente es invariante respecto a las permutaciones de las filas y columnas en la matriz A ; la multiplicación de los elementos de cualquier fila de la matriz A por un escalar a sustituye per A por a per A ; el desarrollo de Laplace en menores respecto a las filas es válido también para los permanentes.

Una lista $(a_{1\sigma(1)}, a_{2\sigma(2)}, \dots, a_{m\sigma(m)})$ de elementos de la matriz A , en la cual ningún par se encuentra en una línea (fila o columna) se denominará *transversal*. El permanente de la matriz A representa, de este modo, una suma de todas las $n!(n - m)$ transversales de los productos de los elementos de la matriz A , dispuestos en una misma transversal.

La determinación de permanente y de sus propiedades señalan que la noción de permanente es próxima a la de determinante. Sin embargo, si la técnica de cálculo de los determinantes está actualmente bien elaborada, el cálculo de los permanentes es un problema técnicamente difícil, cuya resolución requiere nuevos avances. Hoy día sabemos sólo un método que consiste en la aplicación de la fórmula de inclusiones y exclusiones.

Denotemos con A_r una matriz que se obtiene de la matriz A reemplazando r de sus columnas por columnas compuestas de ceros. Sea $S(A_r)$ el producto de las sumas de las filas A_r , y $\sum S(A_r)$, las sumas $S(A_r)$ respecto de todas las selecciones para A_r . En este caso es válida la siguiente fórmula de Ryser:

$$\text{per } A = \sum_{i=0}^{m-1} (-1)^i \binom{n-m+i}{i} \sum S(A_{n-m+i}).$$

En efecto, designemos con M el conjunto de todas las m -permutaciones $(j_1, \dots, j_m)$ con repeticiones de los números naturales $1, 2, \dots, n$. Supongamos que el peso de tal permutación es igual a $\prod_{i=1}^m a_{ij_i}$, y la propiedad p_i

significa que dicha permutación no contiene el número entero i ($i = 1, 2, \dots, n$). Supongamos ahora que A_r se obtuvo de A por sustitución de las columnas con los números $i_1, i_2, \dots, i_r$ por las compuestas de ceros. Entonces, la suma de los pesos de las permutaciones de M que poseen cada una de las propiedades $p_{i_1}, p_{i_2}, \dots, p_{i_r}$ es igual a

$$V(p_{i_1}, p_{i_2}, \dots, p_{i_r}) = S(A_r).$$

Por consiguiente,

$$V(r) = \sum_{1 \leq i_1 < i_2 < \dots < i_r \leq n} V(p_{i_1}, p_{i_2}, \dots, p_{i_r}) = \sum S(A_r).$$

La función per A es igual a la suma de los pesos de los elementos de M que satisfacen las $n - m$ propiedades p_i ($i = 1, \dots, n$). Al hacer uso de la fórmula del método de inclusiones y exclusiones, llegamos a la fórmula de Ryser aducida más arriba. Para una matriz cuadrada A de orden n esta fórmula adopta la forma

$$\text{per } A = \sum_{i=0}^{n-1} (-1)^i \sum S(A_i).$$

Si J es una matriz de orden n , compuesta íntegramente por unidades, entonces per $J = n!$, y a título de corolario, obtenemos la identidad

$$n! = \sum_{i=0}^{n-1} (-1)^i \binom{n}{i} (n-i)^n.$$

Si consideramos la matriz $J - I$, donde I es la matriz unidad, tenemos

$$\text{per } (J - I) = n! \sum_{i=0}^n (-1)^i \frac{1}{i!} = D_n,$$

donde D_n es el número de desórdenes. Valiéndonos de la fórmula de Ryser, podemos obtener otra fórmula para el número de desórdenes:

$$D_n = \sum_{i=0}^{n-1} (-1)^i \binom{n}{i} (n-i)^i (n-i-1)^{n-1}.$$

Existe una conexión interesante entre el permanente de una matriz cuadrada $A = [a_{ij}]$ ($i, j = 1, 2, \dots, n$) de orden n y los permanentes y determinantes de sus submatrices principales. Designemos con $Q_{k,n}$ el conjunto de todas las sucesiones crecientes $\bar{j} = (j_1, \dots, j_k)$ de longitud k de los números naturales $1, 2, \dots, n$: $1 \leq j_1 < j_2 < \dots < j_k \leq n$. Sea $\bar{j} \in Q_{k,n}$; con $A[\bar{j}]$ denotemos una submatriz $[a_{j_s j_t}]$ ($s, t = 1, 2, \dots, k$) de la matriz A , y con $A(\bar{j})$, una submatriz de la matriz A de orden $(n - k)$, complementaria a $A[\bar{j}]$.

Resulta válida la siguiente relación (véase [66]):

$$\text{per } A = \sum_{k=1}^n (-1)^{k-1} \sum_{\bar{j} \in Q_{k,n}} \text{per } A(\bar{j}) \det A[\bar{j}]. \quad (1)$$

Efectivamente, no es difícil notar que la expresión en el segundo miembro de esta fórmula representa una suma de los productos de los elementos de la matriz A dispuestos en una misma transversal con ciertos coeficientes enteros. Por consiguiente,

$$\sum_{k=1}^n (-1)^{k-1} \sum_{\bar{j} \in Q_{k,n}} \text{per } A(\bar{j}) \det A[\bar{j}] = \sum_{\sigma \in S_n} d_\sigma \prod_{i=1}^n a_{i\sigma(i)},$$

donde la sumación se realiza respecto de todas las sustituciones σ del grupo simétrico $\mathfrak{S}_n$ de grado n . Demostremos que $d_\sigma = 1$ para cualquier sustitución $\sigma \in \mathfrak{S}_n$. Con ello quedará demostrada la fórmula (1). Desarrollemos σ en producto de ciclos: $\sigma = (c_1) \dots (c_m)$, y notemos que la expresión $a_{1\sigma(1)} \dots a_{n\sigma(n)}$ figura en aquellos y sólo aquellos productos per $A(j)$ det $A[j]$ del segundo miembro de la fórmula (1), para los cuales existen los números naturales $l_1, \dots, l_i$ ($1 \leq l_1 \leq \dots < l_i \leq m$, $1 \leq i \leq m$) tales que $\{j_1, \dots, j_k\} = \bigcup_{p=1}^m \tilde{c}_p$, donde $\tilde{c}_p$ es el conjunto de números naturales que forman la notación del l_p -ésimo ciclo. Por consiguiente

$$\begin{aligned} d_\sigma &= \sum_{i=1}^m \sum_{j \in Q_{i,m}} (-1)^{|l_1| \dots |l_i| \geq -1} \cdot (-1)^{|l_1| \dots |l_i| \geq -1} \dots \\ &= \sum_{i=1}^m \sum_{j \in Q_{i,m}} (-1)^{i+1} = \sum_{i=1}^m \binom{m}{i} (-1)^{i+1} = 1, \end{aligned}$$

donde $|l_1|, \dots, |l_i|$ es la longitud sumaria de los ciclos $c_{l_1}, \dots, c_{l_i}$. La afirmación está demostrada.

Observemos que la identidad (1) puede ser escrita en la siguiente forma

$$\det A = \sum_{k=1}^n (-1)^{k-1} \sum_{j \in Q_{k,n}} \det A[j] \text{ per } A[j].$$

He aquí dos identidades de MacMahon [72] que pueden obtenerse como simples consecuencias de la fórmula (1):

$$\text{per } A = \sum_{k=1}^n (-1)^{n-k} k! \sum_{N/N_1, \dots, N_k} \prod_{i=1}^k \det A[N_i], \quad (3)$$

$$\det A = \sum_{k=1}^n (-1)^{n-k} k! \sum_{N/N_1, \dots, N_k} \prod_{i=1}^k \text{per } A[N_i]. \quad (4)$$

La suma se toma aquí respecto de todas las particiones del conjunto $N = \{1, 2, \dots, n\}$ en k partes $N_1, \dots, N_k$. En efecto, sustituyendo per $A(j)$ en el segundo miembro de la fórmula (1) por la expresión correspondiente según la misma fórmula y haciendo así un número finito de veces, llegamos, tras algunas transformaciones sencillas, a (3). La identidad (4) se demuestra de un modo análogo empleando la fórmula (2). Las dificultades que existen con los cálculos de los permanentes condujeron a la aparición de desigualdades, las cuales desempeñaron un papel fundamental en la investigación de toda una serie de problemas del análisis combinatorio: problemas de Van der Waerden sobre el mínimo del permanente de una matriz dos veces estocástica, problemas de estimación del número de ternas de Steiner no isomorfas y de rectángulos latinos, problemas sobre el número de recubrimientos de un grafo dimensional reticular por dímeros, etc.

Cronológicamente la primera desigualdad no trivial para los permanentes fue obtenida en 1961 por Marcus y Newman [18]. Sea A una $(m \times n)$ -matriz, y sea B una $(n \times m)$ -matriz. Entonces

$$|\text{per}(A \cdot B)|^2 \leq \text{per}(A \cdot A^T) \cdot \text{per}(B \cdot B^T),$$

con la particularidad de que la igualdad se consigue, si y sólo si existe una fila nula en A , o bien una columna nula en B , o bien $A = DPB^T$ para cierta matriz diagonal D y la matriz conmutadora P . El signo T denota la operación de transposición de la matriz. Para una matriz cuadrada de orden n de aquí se deduce (si ponemos $B = I$):

$$|\text{per} A|^2 \leq \text{per}(A \cdot A^T).$$

La igualdad se verifica, si y sólo si A tiene una fila nula o bien $A = DP$.

En 1980 Egórychev [67], haciendo uso de la representación del permanente en forma de un discriminante mixto, introdujo en la teoría de permanentes las desigualdades clásicas de Alexándrov y Brunn-Minkowski que se emplean en la geometría. Debido a ello él tuvo éxito en demostrar la validez de la conocida hipótesis de Van der Waerden¹⁾.

Aducimos estas desigualdades para los permanentes, dejando al lado su fuente originaria, es decir, la teoría geométrica de los discriminantes mixtos de A. D. Alexándrov, desarrollada por él en los años 30 y utilizada para estudiar los volúmenes de los cuerpos convexos en el espacio euclidiano.

Nos harán falta algunos hechos del álgebra lineal. Sea L un espacio lineal de dimensión finita sobre $\mathbb{R}$, y sea φ una forma bilineal simétrica en L . Suele decirse que (L, φ) es un *espacio con métrica*. Se denomina *signatura* de φ a la terna (k_1, k_2, k_3) , donde k_1 es el número de valores propios positivos de una matriz de la forma φ en cierta base, k_2 , el número de valores propios negativos, y k_3 , el número de valores propios nulos de la matriz mencionada. El teorema de Sylvester (la ley de inercia) afirma que dicha terna no depende de la base. Se dice que el espacio con métrica (L, φ) es el *espacio de Minkowski*, si la signatura de φ es igual a $(1, n-1, 0)$.

Lema 1. Un espacio con métrica (L, φ) es el de Minkowski, si y sólo si se cumple la condición Φ : existe un vector $a \in L$ tal que $\varphi(a, a) > 0$, y para cualquier $b \in L$ de $\varphi(a, b) = 0$, $b \neq 0$, se deduce que $\varphi(b, b) < 0$.

Demostración. Recordemos que el vector $a \in L$ se denomina regular, si existe tal $b \in L$ que $\varphi(a, b) \neq 0$. Un subespacio $M \subseteq L$ es regular, si lo es cualquier vector no nulo suyo. Supongamos ahora que (L, φ) es un espacio con métrica, para el cual se cumple la condición Φ . Por cuanto el vector a es regular, podemos descomponer $L = \{a\} \oplus L_a$, donde $\{a\}$ es la cápsula lineal de a , y L_a el complemento ortogonal de a (véase [69]). Cons-

¹⁾ D. I. Falikman [68] obtuvo independientemente la solución del problema de Van der Waerden (sin demostrar la unicidad de la matriz minimizadora).

truyamos en L_a una base ortogonal $b_2, \dots, b_n$. Entonces $a, b_2, \dots, b_n$ será la base ortogonal en L , donde $\varphi(a, a) > 0$, y $\varphi(b_i, b_i) < 0$, cualquiera que sea $i \in \{2, \dots, n\}$, es decir, (L, φ) es el espacio de Minkowski. Lo recíproco, evidentemente, también tiene lugar (partimos del hecho de que en un espacio con métrica existe una base ortogonal). El lema está demostrado.

Sea (L, φ) un espacio con métrica. Un vector $a \in L$ se llama *positivo* (*negativo*), si $\varphi(a, a) > 0$ ($\varphi(a, a) < 0$), e *isótropo*, si $\varphi(a, a) = 0$.

Teorema. (Desigualdad de Cauchy—Buniakovski para el espacio de Minkowski). Sea (L, φ) un espacio de Minkowski; a , un vector positivo y b , un vector arbitrario. Entonces

$$\varphi^2(a, b) \geq \varphi(a, a) \cdot \varphi(b, b) \quad (5)$$

y la igualdad tiene lugar, cuando y sólo cuando $b = \lambda a$ para cierto $\lambda \in \mathbb{R}$.

Demostración. Si $b = \lambda a$ para cierto $\lambda \in \mathbb{R}$, entonces en (5) tiene lugar la igualdad. Si $\varphi(b, b) \leq 0$, entonces, evidentemente, (5) se verifica. Supongamos que $b \notin \{a\}$, $\varphi(b, b) > 0$; en este caso examinemos un plano $P = [a, b]$. Demostraremos que en P hay un vector negativo. Si para cierto $y \in P$, $y \neq 0$, se verifica $\varphi(y, y) = 0$, entonces, según el lema 1, tenemos $\varphi(a, y) \neq 0$. De aquí resulta que la expresión $\varphi(a + \lambda y, a + \lambda y) = \varphi(a, a) + 2\lambda\varphi(a, y)$ será inferior a cero con λ correspondiente, es decir, se ha encontrado un vector negativo en P . Supongamos ahora que para todo $x \in P$, $x \neq 0$, tiene lugar $\varphi(x, x) > 0$. En este caso P es un subespacio regular. Por eso se construye una descomposición en la suma directa: $L = P \oplus P_1$, donde P_1 es un complemento ortogonal de P . Tomemos las bases ortogonales; $\{e_1, e_2\}$ en P y $\{e_3, \dots, e_n\}$ en P_1 . La unión de ellas es una base ortogonal para L . En esta base $\varphi(e_1, e_1) > 0$, $\varphi(e_2, e_2) > 0$, por consiguiente, la signatura de L es (k_1, k_2, k_3) con $k_1 \geq 2$, lo que contradice la afirmación de que (L, φ) es un espacio de Minkowski. Así pues, la suposición de partida no es cierta y en P siempre existe un vector negativo. Examinemos ahora la expresión $\varphi(a + \lambda b, a + \lambda b) = \varphi(a, a) + 2\lambda\varphi(a, b) + \lambda^2\varphi(b, b)$; para cierto λ ella es negativa, por lo cual el trinomio cuadrado de λ cuenta con raíces reales. Por consiguiente, su discriminante no es negativo:

$$\varphi^2(a, b) - \varphi(a, a) \cdot \varphi(b, b) \geq 0,$$

lo que demuestra (5). El caso de igualdad se analiza con ayuda de los razonamientos análogos y se sugiere que el lector lo haga como un ejercicio. El teorema queda demostrado.

Teorema (desigualdad de Alexándrov-Egórychev para permanentes). Sea $a_{ij} \geq 0$ para $i, j = 1, \dots, n$. Entonces se verifica la desigualdad

$$\text{per}^2 \left[\begin{array}{cccc} a_{1,1} & \dots & a_{1,n-2} & a_{1,n-1} & a_{1,n} \\ \dots & \dots & \dots & \dots & \dots \\ a_{n,1} & \dots & a_{n,n-2} & a_{n,n-1} & a_{n,n} \end{array} \right] \geq$$

$$\geq \left[\frac{a_{1,1} \dots a_{1,n-2} a_{1,n-1} a_{1,n-1}}{a_{n,1} \dots a_{n,n-2} a_{n,n-1} a_{n,n-1}} \right] \cdot \text{per} \left[\frac{a_{1,1} \dots a_{1,n-2} a_{1,n} a_{1,n}}{a_{n,1} \dots a_{n,n-2} a_{n,n} a_{n,n}} \right]. \quad (6)$$

Las matrices que están en el segundo miembro difieren sólo por las dos últimas columnas: en una matriz la última columna repite la penúltima, en la otra, viceversa. Si $a_{ij} > 0$ para $i = 1, \dots, n$; $j = 1, \dots, n-2$, entonces la igualdad se consigue, cuando y sólo cuando para todo $i = 1, 2, \dots, n$ tiene lugar $a_{i,n-1} = \lambda a_{i,n}$, donde $\lambda \in \mathbb{R}$. Para demostrar el teorema se requieren dos lemas de D. I. Falikman [68].

Lema 2. Sea $\varphi(x, y)$ una función bilineal simétrica en el espacio vectorial V tal, que existe un vector no nulo $a \in V$, el cual posee la propiedad de que $\varphi(a, a) = 0$, y para todo $x \in V$, $x \neq \lambda a$ ($\lambda \in \mathbb{R}$) se afirma que de la igualdad $\varphi(x, a) = 0$ se deduce que $\varphi(x, x) < 0$. Entonces, si $b, c \in V$, $b, c \neq 0$, y $\varphi(b, c) = 0$, $\varphi(b, b) > 0$, entonces $\varphi(c, c) < 0$.

Demostración. Si para cierto $x \in V$ se cumple la condición $\varphi(x, a) = 0$, entonces $\varphi(x, x) \leq 0$. De la desigualdad $\varphi(b, b) > 0$ se deduce $\varphi(b, a) \neq 0$. Pongamos $\eta = -\varphi^{-1}(b, a) \cdot \varphi(c, a)$. Entonces $\varphi(c + \eta b, a) = 0$, y, por consiguiente,

$$0 \geq \varphi(c + \eta b, c + \eta b) = \varphi(c, c) + \eta^2 \varphi(b, b).$$

Si $\eta \neq 0$, entonces $\varphi(c, c) \leq -\eta^2 \varphi(b, b) < 0$. En cambio, si $\eta = 0$, tenemos $\varphi(c, a) = 0$. Los vectores c y a no son colineales, puesto que de lo contrario de la condición $\varphi(b, c) = 0$ se desprendería que $\varphi(b, a) = 0$, lo que no es cierto. Por eso, $\varphi(c, c) < 0$ también en este caso.

Lema 3. Sea $\varphi(x, y) = \text{per} \begin{bmatrix} a_{1,1} & \dots & a_{1,n-2} x_1 y_1 \\ \dots & \dots & \dots & \dots & \dots \\ a_{n,1} & \dots & a_{n,n-2} x_n y_n \end{bmatrix}$, donde $a_{ij} > 0$ para i, j tales que $1 \leq i \leq n$, $1 \leq j \leq n-2$, $n \geq 2$.

Entonces, si $b, c \in \mathbb{R}^n$, $b, c \neq 0$, $\varphi(b, c) = 0$, $\varphi(b, b) > 0$, se tiene que $\varphi(c, c) < 0$.

La demostración del lema la realizaremos por inducción respecto de $n \geq 2$. Cuando $n = 2$, $\varphi(x, y) = \text{per} \begin{bmatrix} x_1 y_1 \\ x_2 y_2 \end{bmatrix} = x_1 y_2 + x_2 y_1$. Tenemos: $0 = \varphi(b, c) = b_1 c_2 + b_2 c_1$, $\varphi(b, b) = 2b_1 b_2 > 0$. Por consiguiente, $b_1 \neq 0$, $b_2 \neq 0$ y $c_1 \neq 0$, $c_2 \neq 0$. Al multiplicar ambos miembros de la igualdad $-b_1 c_2 = b_2 c_1$ por $b_1 c_2$, obtenemos $-b_1^2 c_2^2 = b_1 b_2 c_1 c_2$. Por consiguiente, $c_1 c_2 < 0$, y $\varphi(c, c) = 2c_1 c_2 < 0$.

Supongamos ahora que $n > 2$, y que para todos los k naturales: $2 \leq k \leq n-1$ y la afirmación del lema es válida. Pongamos $a = (0, \dots, 0, 1)$. Entonces

$$\varphi(a, a) = \text{per} \begin{bmatrix} a_{1,1} & \dots & a_{1,n-2} & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ a_{n-1,1} & \dots & a_{n-1,n-2} & 0 & 0 \\ a_{n,1} & \dots & a_{n,n-2} & 1 & 1 \end{bmatrix} = 0.$$

Elijamos un vector arbitrario $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ tal que $\varphi(x, a) = 0$, y los vectores x y a no son colineales, es decir, $x' = (x_1, \dots, x_{n-1}) \neq 0$. Entonces $\varphi(x, x) < 0$. En efecto,

$$\begin{aligned} \varphi(x, a) &= \text{per} \begin{bmatrix} a_{1,1} & \dots & a_{1,n-2} & x_1 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ a_{n,1} & \dots & a_{n,n-2} & x_n & 1 \end{bmatrix} = \\ &= \text{per} \begin{bmatrix} a_{1,1} & \dots & a_{1,n-2} & x_1 & \\ \dots & \dots & \dots & \dots & \\ a_{n-1,1} & \dots & a_{n-1,n-2} & x_{n-1} & \end{bmatrix} = 0 \end{aligned}$$

Por consiguiente

$$\varphi(x, x) = \text{per} \begin{bmatrix} a_{1,1} & \dots & a_{1,n-2} & x_1 & x_1 \\ \dots & \dots & \dots & \dots & \dots \\ a_{n,1} & \dots & a_{n,n-2} & x_n & x_n \end{bmatrix} = \sum_{i=1}^{n-2} a_{ni} \varphi_i(x', x'),$$

donde

$$\begin{aligned} \varphi_i(x', x') &= \\ &= \text{per} \begin{bmatrix} a_{1,1} & \dots & a_{1,i-1} & a_{1,i+1} & \dots & a_{1,n-2} & x_1 & x_1 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{n-1,1} & \dots & a_{n-1,i-1} & a_{n-1,i+1} & \dots & a_{n-1,n-2} & x_{n-1} & x_{n-1} \end{bmatrix} \end{aligned}$$

Está claro que $\varphi_i(x', a'_i) = \varphi(x, a) = 0$, y, además, $\varphi_i(a'_i, a'_i) > 0$ (puesto que $a_{ij} > 0$ para $i = 1, \dots, n, j = 1, \dots, n-2$), donde $a'_i = (a_{1,i}, \dots, a_{n-1,i})$. Por consiguiente, de acuerdo con la hipótesis de inducción, $\varphi_i(x', x') < 0$, de donde $\varphi(x, x) < 0$. La aplicación del lema anterior finaliza la demostración del lema 3.

Demostremos ahora (6). Supongamos al principio que todos los elementos de la matriz $A = \|a_{ij}\|$ son positivos. Para la función φ , construida en el lema 3, existe un vector positivo $a = (1, \dots, 1)$. De acuerdo con el lema 3, de que $\varphi(a, b) = 0, b \neq 0$, se desprende $\varphi(b, b) < 0$. Por eso, según el lema 1, el par $(\mathbb{R}^n, \varphi)$ es un espacio de Minkowski, y la desigualdad (6) es la desigualdad de Cauchy—Buniakovski (5) para dicho espacio.

El caso en que entre los elementos a_{ij} pueden encontrarse elementos nulos se desprende del examinado más arriba con ayuda de un paso límite. Sin embargo, en estas circunstancias no es obligatorio que el signo de igualdad tenga lugar sólo cuando los vectores b y c son colineales. El caso general en que (6) se convierte en una igualdad ha sido analizado por A. A. Panov [70].

En 1926 Van der Waerden [71] planteó una cuestión; ¿cuál será el valor mínimo de la función del permanente sobre el conjunto Ω_n de todas las matrices dos veces estocásticas de orden n ? La respuesta positiva a esta pregunta era conocida como hipótesis de Van der Waerden. Se suponía que si A era una matriz dos veces estocástica de orden n , entonces per

$A \geq n^{-n} \cdot n!$, con la particularidad de que la igualdad se consigue, cuando y sólo cuando A es la matriz J_n , todos los elementos de la cual son iguales a n^{-1} . La demostración de la validez de la hipótesis ha sido obtenida en 1980. Comencemos su exposición por una serie de lemas auxiliares (véase [18]). Introduzcamos algunas nociones preliminares.

Una matriz no negativa de orden n la llamaremos *parcialmente descomponible* si contiene una submatriz nula de dimensión $k \times (n - k)$. En otras palabras, la matriz A es parcialmente descomponible, si existen matrices conmutables P y Q de tal índole que

$$PAQ = \begin{bmatrix} B & C \\ 0 & D \end{bmatrix},$$

donde B y D son matrices cuadradas. Si la matriz no contiene submatriz nula de dimensión $k \times (n - k)$ para ningún $k = 1, \dots, n - 1$, se denominará *bien indescomponible*.

Observemos que si A es una matriz dos veces estocástica parcialmente descomponible, se encontrarán matrices conmutables P y Q tales que PAQ se descomponga en una suma directa de las matrices dos veces estocástica.

En efecto, supongamos que

$$PAQ = \begin{bmatrix} B & C \\ 0 & D \end{bmatrix},$$

donde B y D son matrices cuadradas, y el orden D es igual a k . Denotemos con $s(X)$ la suma de los elementos en la matriz X . Puesto que PAQ es una matriz dos veces estocástica, la suma de los elementos dispuestos en sus primeras $n - k$ columnas, es igual a $n - k$, es decir, $s(B) = n - k$. Análogamente, $s(D) = k$. Por consiguiente, $s(C) = s(PAQ) - s(B) - s(D) = n - (n - k) - k = 0$. Por cuanto C es no negativa, tenemos $C = 0$. Es obvio que las matrices B y D son dos veces estocásticas.

Lema 4. Sea A una matriz de orden n . Para que cada transversal suya contenga cero, es necesario y suficiente que en A exista una submatriz nula de dimensión $r \times t$, donde $r + t = n + 1$.

La validez de la afirmación se deduce del teorema de König (véase § 4.1).

Denotemos con A_{ij} una submatriz de la matriz A obtenida por supresión de la i -ésima fila y de la j -ésima columna.

Lema 5. Una matriz no negativa A de orden $n \geq 2$ es bien indescomponible, cuando y sólo cuando para todos los $i, j = 1, \dots, n$ se verifica la desigualdad: per $A_{ij} > 0$.

Demostración. Según el lema 4 per $A_{ij} = 0$, cuando y solo cuando la submatriz A_{ij} , y, por tanto, la matriz A contienen una submatriz nula de dimensión $r \times t$, donde $r + t = (n - 1) + 1$. Dicho de otro modo, per $A_{ij} = 0$, si y sólo si A es parcialmente descomponible.

La matriz dos veces estocástica A de orden n se llamará *minimizadora*, siempre que per $A = \min_{B \in \Omega} \text{per } B$.

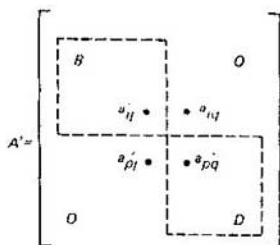


Fig. 4.2.

Lema 6. La matriz minimizadora es bien indecomponible.

Demostración. Sea $A \in \Omega_n$ una matriz minimizadora y supongamos que A es parcialmente descomponible. En este caso existirán matrices conmutables P y Q tales que PAQ será una suma directa de las matrices $B \in \Omega_{n-k}$ y $D \in \Omega_k$. Denotemos $A' = PAQ$. Está claro que $A' \in \Omega_n$, A' es una matriz minimizadora. Elijamos en la transversal de la matriz A' , compuesta por los elementos no nulos, los elementos a'_{ij} y a'_{pq} que pertenecen a las matrices B y D , respectivamente. Tenemos: $a'_{ij} \neq 0$, $a'_{pq} \neq 0$, $a'_{iq} = a'_{pj} = 0$, per $A'_{ij} > 0$ y per $A'_{pq} > 0$. Del lema 4 se deduce que per $A'_{iq} = \text{per } A'_{pj} = 0$. Veamos la matriz $A'(\varepsilon)$: $a'_{ij}(\varepsilon) = a'_{ij} - \varepsilon$, $a'_{pq}(\varepsilon) = a'_{pq} - \varepsilon$, $a'_{iq}(\varepsilon)a'_{pj}(\varepsilon) = \varepsilon$, $a'_{it}(\varepsilon) = a'_{it}$, si $r \neq i, p$, o bien $t \neq j, q$. Está claro que $A'(\varepsilon) \in \Omega_n$, cuando los valores de $\varepsilon > 0$ son suficientemente pequeños. Nos convencemos inmediatamente de que $\frac{d}{d\varepsilon} \text{per } A'(\varepsilon)|_{\varepsilon=0} = -\text{per } A'_{ij} + \text{per } A'_{iq} - \text{per } A'_{pq} + \text{per } A'_{ij} = -\text{per } A'_{ij} - \text{per } A'_{pq} < 0$. Por consiguiente, per $A'(\varepsilon) < \text{per } A'$, cuando los valores de $\varepsilon > 0$ son suficientemente pequeños. Esto contradice el hecho de que la matriz A' es minimizadora.

Lema 7. Si una matriz $A \in \Omega_n$ es minimizadora y si $a_{pq} > 0$, entonces per $A_{pq} = \text{per } A$.

Demostración. Estudiemos el siguiente conjunto de las matrices de dimensión $n \times n$: $C(A) = \{X = [x_{ij}] | x_{ij} \geq 0 \text{ para cualesquiera } i, j = 1, \dots, n; x_{ij} = 0 \text{ para } (i, j) \in Z\}$, donde $Z = \{(i, j) | a_{ij} = 0\}$. Consideramos $C(A)$ como un conjunto de puntos en un espacio euclidiano de dimensión $n^2 - |Z|$; en este caso A será un punto interior de dicho conjunto. En $C(A)$ tenemos un problema suave (véase [11]):

$$\text{per } X \rightarrow \inf$$

con restricciones

$$\sum_{j=1}^n x_{ij} = 1, \quad i = 1, \dots, n;$$

$$\sum_{i=1}^n x_{ij} = 1, \quad j = 1, \dots, n-1 \quad (7)$$

(observemos que la condición $\sum x_{in} = 1$ es una consecuencia de las condiciones (7)). La matriz A es la solución de este problema, por eso existen los multiplicadores de Lagrange $\lambda_0, \lambda_1, \dots, \lambda_n, \mu_1, \dots, \mu_{n-1}$, que no son nulos simultáneamente, tales que la función de Lagrange

$$F(X, \lambda_0, \lambda_1, \dots, \lambda_n, \mu_1, \dots, \mu_{n-1}) = \lambda_0 \text{ per } X - \sum_{i=1}^n \lambda_i \left(\sum_{k=1}^n x_{ik} - 1 \right) - \sum_{j=1}^{n-1} \mu_j \left(\sum_{k=1}^n x_{kj} - 1 \right)$$

satisface en el punto $X = A$ las condiciones de carácter estacionario respecto de x (véase [111, p. 47—48]):

$\frac{\partial}{\partial x_{ij}} F(X, \lambda_0, \lambda_1, \dots, \lambda_n, \mu_1, \dots, \mu_{n-1})|_{X=A} = 0$ para cualesquiera $(i, j) \notin Z$, es decir, tenemos:

$$\begin{cases} \lambda_0 \text{ per } A_{ij} = \lambda_i + \mu_j \text{ para } i = 1, \dots, n; j = 1, \dots, n-1; \\ (i, j) \notin Z, \\ \lambda_0 \text{ per } A_{in} = \lambda_i \text{ para } i = 1, \dots, n; (i, n) \notin Z. \end{cases} \quad (8)$$

Supongamos que $\lambda_0 = 0$. Entonces (8) se reduce a

$$\begin{cases} \lambda_i + \mu_j = 0; i = 1, \dots, n; j = 1, \dots, n-1; (i, j) \notin Z, \\ \lambda_i = 0; i = 1, \dots, n; (i, n) \notin Z. \end{cases} \quad (9)$$

De la existencia (i, n) tales que $a_{in} \neq 0$, se desprende que se tienen i , para los cuales $\lambda_i = 0$. Sea $\lambda_1 = \dots = \lambda_k = 0$. Si $k = n$, entonces, de acuerdo con (9), de la existencia de $(i, j) \notin Z$, para cada $j = 1, \dots, n-1$ obtenemos: $\mu_1 = \dots = \mu_{n-1} = 0$, que contradice la afirmación de que no todos los multiplicadores de Lagrange son nulos. Supongamos que $k < n$ y $\lambda_{k+1} \neq 0, \dots, \lambda_n \neq 0$ (esto puede conseguirse por permutación de filas en A). Si existe un elemento no nulo a_{ij} , donde $i \in \{1, \dots, k\}, j < n$, entonces de (9) obtenemos: $\mu_j = 0$. Cambiemos de lugar las columnas de A de un modo tal que se cumplan las condiciones: $\mu_1 = \dots = \mu_{n-1-l} = 0$, y $\mu_j \neq 0$, cualesquiera que sean $n-1-l < j < n$. Si $l = 0$, entonces $\mu_1 = \dots = \mu_{n-1} = 0$, y, por consiguiente, $\lambda_{k+1} = 0$, puesto que existe $(k+1, j) \notin Z$, donde $j < n$ (otra vez aplicamos (9)). Esto contradice la suposición de que $\lambda_{k+1} \neq 0$. Sea $l > 0$, entonces una submatriz de la matriz A , formada por sus primeras filas y columnas con los números $n-l, \dots, n-1$ es nula y tiene las dimensiones $k \times l$. Por cuanto A es bien indecomponible, entonces, según el lema 5, $\text{per } A_{n,n} > 0$, y de aquí obtenemos con arreglo al lema 4:

$$k + l < n. \quad (10)$$

Demostremos ahora que la submatriz A' , formada por las filas $k+1, \dots, n$ y las columnas $1, \dots, n-l-1$ es nula. En efecto, si en dicha submatriz existe un elemento no nulo a_{ij} , entonces, debido a que $\mu_j = 0$, y en virtud de la igualdad $\lambda_i + \mu_j = 0$ (de (9)) obtenemos $\lambda_i = 0$, pero esto no es posible

por construcción. Esta submatriz tiene las dimensiones $(n - k) \times (n - l - 1)$. De aquí y de la desigualdad (10): $n - k + n - l - 1 = n + (n - k - l - 1) \geq n$, puesto que $n - k - l - 1 \geq 0$. Lo último contradice la plena descomponibilidad de la matriz A . Esto quiere decir que la igualdad $\lambda_0 = 0$ es imposible. Al dividir las relaciones (8) por λ_0 , obtenemos

$$\text{per } A_{ij} = \tilde{\lambda}_i + \tilde{\mu}_i \text{ para cualesquiera } (i, j) \notin Z \\ i, j \in \{1, \dots, n\}$$

(aquí tomamos $\tilde{\mu}_n = 0$).

Permutando, si es necesario, las filas y columnas de la matriz A , logremos que se cumplan las desigualdades: $\tilde{\lambda}_1 \leq \tilde{\lambda}_2 \leq \dots \leq \tilde{\lambda}_n$, $\tilde{\mu}_1 \leq \tilde{\mu}_2 \leq \dots \leq \tilde{\mu}_n$. Descomponiendo per A respecto de la primera fila y la última columna, y empleando el hecho de que la suma de los elementos en cualquier fila (columna) de la matriz A es igual a la unidad, obtendremos

$$\lambda_1 + \tilde{\mu}_n \geq \sum_{\substack{j=1 \\ (1,j) \notin Z}}^n a_{1j} \text{ per } A_{1j} = \sum_{j=1}^n a_{1j} \text{ per } A_{1j} \text{ per } A_{1j} = \text{per } A = \\ = \sum_{i=1}^n a_{in} \text{ per } A_{in} = \sum_{\substack{i=1 \\ (i,n) \notin Z}}^n a_{in} \text{ per } A_{in} \geq \tilde{\lambda}_1 + \tilde{\mu}_n.$$

Por consiguiente, $\tilde{\lambda}_1 + \tilde{\mu}_n = \text{per } A = \text{per } A_{1j} = \text{per } A_{in}$ para todos los i, j tales que $(1, j) \notin Z$, $(i, n) \notin Z$.

Supongamos que para todos los i, j tales que $j \geq j_1$, $i \leq i_1$, $(i, j) \notin Z$ ya se ha demostrado que $\text{per } A_{ij} = \text{per } A$. Veamos la $(i_1 + 1)$ -ésima fila de la matriz A . Si existe $j \geq j_1$ tal que $(i_1 + 1, j) \notin Z$, entonces para todo $j \geq j_1$ tal que $(i_1 + 1, j) \notin Z$, tenemos: $\text{per } A_{i_1+1,j} = \text{per } A$ (por suposición) y para todo $j < j_1$ tal que $(i_1 + 1, j) \notin Z$ tenemos: $\text{per } A_{i_1+1,j} = \tilde{\lambda}_{i_1+1} + \tilde{\mu}_j \leq \tilde{\lambda}_{i_1+1} + \tilde{\mu}_{j_1} = \text{per } A_{i_1+1,j_1} = \text{per } A$. Por consiguiente, $\text{per } A = \sum_{j=1}^n a_{i_1+1,j} \text{ per } A_{i_1+1,j} \leq \text{per } A$, y para todos los j tales que $(i_1 + 1, j) \notin Z$, se verifica la igualdad

$$\text{per } A_{i_1+1,j} = \text{per } A.$$

Los razonamientos análogos son lícitos también para la $(j_1 - 1)$ -ésima columna de la matriz A . En cambio, si para cualesquiera $j \geq j_1$ y $i \leq i_1$ se verifica:

$$(i_1 + 1, j) \in Z \text{ y } (i, j_1 - 1) \in Z, \text{ es decir, } a_{i_1+1,j} = 0 \text{ y } a_{i,j_1-1} = 0,$$

entonces

$$\tilde{\lambda}_{i_1+1} + \tilde{\mu}_{j_1-1} \geq \sum_{j=1}^{j_1-1} a_{i_1+1,j} \text{ per } A_{i_1+1,j} = \text{per } A = \\ = \sum_{i=i_1+1}^n a_{i,j_1-1} \text{ per } A_{i,j_1-1} \geq \tilde{\lambda}_{i_1+1} + \tilde{\mu}_{j_1-1},$$

y, por consiguiente, para cualesquiera i, j tales que $(i, j_1 - 1) \notin Z$, $(i_1 + 1, j) \notin Z$, tenemos: $\text{per } A_{i, j_1 - 1} = \text{per } A_{i_1 + 1, j} = \text{per } A$.

Así pues, si $(i, j) \notin Z$, entonces $\text{per } A_{ij} = \text{per } A$, lo que se trataba de demostrar.

Lema 8 (desigualdad de London). Si A es una matriz minimizadora en Ω_n , para cualesquiera $i, j = 1, \dots, n$ se verifica la desigualdad $\text{per } A_{ij} \geq \text{per } A$.

Demostración. Para toda matriz conmutadora $P = [p_{ij}]$ y el número $\theta \in [0; 1]$ pongamos $f_p(\theta) = \text{per}((1 - \theta)A + \theta P)$. Por cuanto A es la matriz minimizadora, entonces $f'_p(0) \geq 0$ para cualquier matriz conmutadora P . Sea σ una permutación correspondiente a P . Tenemos

$$\begin{aligned} f'_p(0) &= \sum_{r, t=1}^n (-a_{rt} + p_{rt}) \text{per } A_{rt} = \sum_{r, t=1}^n p_{rt} \text{per } A_{rt} - n \text{per } A = \\ &= \sum_{i=1}^n \text{per } A_{r\sigma(r)} - n \text{per } A \geq 0. \end{aligned}$$

Por consiguiente,

$$\sum_{r=1}^n \text{per } A_{r\sigma(r)} \geq n \text{per } A \quad (11)$$

para toda permutación σ . Según el lema 6, la matriz A es bien indescomponible, por lo cual, de conformidad con el lema 5, cualquier elemento de A se dispone en cierta transversal, cuyos elementos restantes son todos positivos. Dicho de otro modo, para cualquier par (i, j) existe tal permutación σ que $j = \sigma(i)$, y $a_{r\sigma(r)} > 0$, cuando $r = 1, \dots, i - 1, i + 1, \dots, n$. Pero, en virtud del lema 7, esto significa que $\text{per } A_{r\sigma(r)} = \text{per } A$ para $r = 1, \dots, i - 1, i + 1, \dots, n$. Por cuanto $j = \sigma(i)$, obtenemos de (11): $\text{per } A_{ij} \geq \text{per } A$, lo que se trataba de demostrar.

Ahora, siguiendo los razonamientos de Egórychev, señalemos cómo se puede demostrar, al hacer uso de la desigualdad (6), la validez de la suposición de Van der Waerden.

Sea A una matriz minimizadora. Probemos que $\text{per } A_{ij} = \text{per } A$ para cualesquiera $i, j = 1, \dots, n$. Supongamos lo contrario: existe al menos un solo par $r, s \in \{1, \dots, n\}$ tal que $\text{per } A_{rs} > \text{per } A$. Por cuanto la matriz A es dos veces estocástica, existe $t \in \{1, \dots, n\}$ tal que $a_{rt} > 0$. Entonces, en virtud de la desigualdad (6) y del lema 8, tenemos

$$\begin{aligned} \text{per}^2 A &\geq \left(\sum_{k=1}^n a_{ks} \text{per } A_{kt} \right) \left(\sum_{k=1}^n a_{kt} \text{per } A_{ks} \right) > \\ &> \left(\sum_{k=1}^s a_{ks} \text{per } A \right) \left(\sum_{k=1}^n a_{kt} \text{per } A \right) = \text{per}^2 A \end{aligned}$$

(el carácter estricto de la desigualdad en este caso se desprende de que $a_{rt} > 0$ y $\text{per } A_{rs} > \text{per } A > 0$).

Demostremos ahora que si A es una matriz minimizadora, entonces $a_{ij} = n^{-1}$ ($i, j = 1, \dots, n$). Supongamos que la columna de la matriz A

de número n se diferencia de $(n^{-1}, \dots, n^{-1})^T$. Por cuanto A no se descompone en una suma directa de matrices no nulas, cada elemento de la n -ésima columna de A es inferior a la unidad. Por consiguiente, entre los primeros $n - 1$ elementos de cada fila de la matriz A hay por lo menos un elemento no nulo. Pasando de A a una matriz que se obtiene a partir de A sustituyendo la i -ésima y j -ésima columnas de su semisuma ($i \neq j$, $i, j = 1, \dots, n - 1$), obtendremos, al hacer un número finito de pasos, la matriz A en la que todos los elementos de las primeras $n - 1$ columnas son positivos, mientras que la n -ésima columna coincide con la n -ésima columna de la matriz A , con la particularidad de que $\text{per } A' = \text{per } A$, puesto que, hecho cualquier paso, el permanente no varía (lo que se desprende de la fórmula de descomposición del permanente de una matriz respecto de las columnas y de la igualdad de todos los permanentes de las submatrices de $(n - 1)$ -ésimo orden).

Designemos con $A_j(b)$ una matriz obtenida de A' como resultado de la sustitución de la j -ésima columna por la columna b . Ya hemos aclarado que en la desigualdad

$$\text{per}^2 A' \geq \text{per } A'_n(a'_i) \text{ per } A'_i(a_n)$$

se realiza la igualdad. El carácter positivo de las componentes $a'_1, \dots, a'_{n-1}$ da el derecho de afirmar que $a_n = \lambda_i a'_i$ ($i = 1, \dots, n - 1$). Por cuanto las sumas de componentes de los vectores son todas iguales a 1, entonces $a_n = a'_i$ ($i = 1, \dots, n - 1$). Como la matriz A' es dos veces estocástica llegamos a que $a'_1 = \dots = a'_{n-1} = a_n = (n^{-1}, \dots, n^{-1})^T$, lo que es una contradicción. Por consiguiente, todas las columnas de la matriz A son iguales a $(n^{-1}, \dots, n^{-1})^T$ y esto demuestra nuestra afirmación.

Veamos el problema sobre las estimaciones superiores para los permanentes. En 1960 Ryser supuso que en la clase de $(0, 1)$ -matrices de orden mk , cada línea de las cuales contiene k unidades, el valor máximo del permanente se logra en la suma directa de matrices de orden k , compuestas por unidades. J. Minc enunció en 1963 una suposición más general de que si A era una $(0, 1)$ -matriz de orden n que tiene $r_1, \dots, r_n$ sumas respecto de las filas, entonces $\text{per } A \leq \prod_{i=1}^n (r_i!)^{1/r_i}$. De esta desigualdad se deduce

inmediatamente la hipótesis de Ryser.

La desigualdad de Minc la demostró por primera vez (en 1973) Bregmán [73]. Más abajo aducimos una demostración más sencilla de esta desigualdad que se debe a Schrijver [74].

Lema 9. Para los números reales no negativos $x_1, \dots, x_r$ es válida la desigualdad

$$\left(r^{-1} \sum_{i=1}^r x_i\right)^{\sum_{i=1}^r x_i} \leq \prod_{i=1}^r x_i^{x_i}$$

(en esta fórmula se supone que $0^0 = 1$).

Demostración. Al hacer uso de la convexidad de la función $y = x \log_2 x$, obtenemos

$$r^{-1} \sum_{i=1}^r x_i \log_2 \left(r^{-1} \sum_{i=1}^r x_i \right) \leq r^{-1} \sum_{i=1}^r x_i \log_2 x_i,$$

de donde se deduce inmediatamente el resultado del lema.

Sea $A = [a_{ij}]$ una $(0, 1)$ -matriz cuadrada de orden n , y sea S un conjunto de todas las permutaciones π de los números $1, \dots, n$ tales que $\prod_{i=1}^n a_{i\pi(i)} = 1$ entonces

$$\prod_{i,j} \text{per}(A_{ij})^{\text{per} A_{ij}} = \prod_{\pi \in S} \prod_{i=1}^n \text{per} A_{i\pi(i)}, \quad \prod_{i=1}^n r_i^{\text{per} A} = \prod_{\pi \in S} \prod_{i=1}^n r_i.$$

La validez de estas identidades se confirma con facilidad mediante una comprobación directa. Aprovechémoslas para demostrar la desigualdad de Minc. Apliquemos el método de inducción matemática completa respecto de n . Del lema 9 tenemos

$$\begin{aligned} (\text{per } A)^{n \text{ per } A} &= \prod_{i=1}^n (\text{per } A)^{\text{per} A_{ij}} = \prod_{i=1}^n \left(\prod_{j=1}^n a_{ij} \text{per } A_{ij} \right)^{\sum_{j=1}^n a_{ij} \text{per} A_{ij}} \leq \\ &\leq \prod_{i=1}^n \left(r_i^{\text{per} A} \prod_{j: a_{ij}=1} (\text{per } A_{ij})^{\text{per} A_{ij}} \right) = \prod_{\pi \in S} \left[\left(\prod_{i=1}^n r_i \right) \left(\prod_{i=1}^n \text{per } A_{i\pi(i)} \right) \right]. \end{aligned}$$

Apliquemos ahora la hipótesis de inducción a cada una de las matrices $A_{i\pi(i)}$:

$$\begin{aligned} \prod_{i=1}^n \text{per } A_{i\pi(i)} &\leq \prod_{i=1}^n \left(\prod_{\substack{j \neq i \\ a_{j\pi(i)}=0}} (r_j!)^{\frac{1}{r_j}} \right) \left(\prod_{\substack{j \neq i \\ a_{j\pi(i)}=1}} (r_j - 1)!^{\frac{1}{(r_j-1)}} \right) = \\ &= \prod_{j=1}^n \left(\prod_{\substack{i \neq j \\ a_{j\pi(i)}=0}} (r_j!)^{\frac{1}{r_j}} \right) \left(\prod_{\substack{i \neq j \\ a_{j\pi(i)}=1}} (r_j - 1)!^{\frac{1}{(r_j-1)}} \right) = \prod_{j=1}^n \left[r_j!^{\frac{n-r_j}{r_j}} (r_j - 1)! \right]. \end{aligned}$$

La última igualdad se ha obtenido por cálculo de los factores $(r_j!)^{1/r_j}$ y $(r_j - 1)!^{1/(r_j-1)}$. Siendo fijos π y j , el número de índices i , para los cuales $i \neq j$, es igual a $n - r_j$ si $a_{j\pi(i)} = 0$, y a $n - r_j - 1$, si $a_{j\pi(i)} = 1$. Así pues,

$$\begin{aligned} (\text{per } A)^{n \text{ per} A} &\leq \prod_{\pi \in S} \left[\left(\prod_{i=1}^n r_i \right) \left(\prod_{j=1}^n r_j!^{\frac{n-r_j}{r_j}} (r_j - 1)! \right) \right] = \\ &= \prod_{\pi \in S} \left(\prod_{i=1}^n r_i!^{\frac{n}{r_i}} \right) = \left(\prod_{i=1}^n r_i! \frac{1}{r_i} \right)^{n \text{ per} A}, \end{aligned}$$

lo que se trataba de demostrar.

Volvamos a las aplicaciones de las desigualdades a los permanentes.

Denotemos con $L(r, n)$ el número de $(r \times n)$ -rectángulos latinos compuestos por los elementos de un conjunto $S = \{1, 2, \dots, n\}$. Analicemos algún rectángulo latino de dimensión $t \times n$, $1 \leq t < r$. De cuántas maneras puede ser extendido hasta que se obtenga un $(t+1) \times n$ -rectángulo latino? Está claro que el número de tales métodos es igual a $\text{per } A$, donde A es la matriz de incidencia del sistema de subconjuntos $(S_1, \dots, S_n)$, donde S_i es un subconjunto de aquellos elementos de S que no se encuentran en la i -ésima columna del rectángulo a extender ($i = 1, \dots, n$).

En cada fila de la matriz A están contenidas $n - t$ unidades; la matriz $(n - t)^{-1}A$ será dos veces estocástica. Aplicando las desigualdades de Minc y de Van der Waerden, obtendremos

$$\frac{n!}{n^n} (n - t)^n \leq \text{per } A \leq (n - t)!^{\frac{n}{n-t}}.$$

Cuando $t = 1$, $\text{per } A = n! \sum_{i=0}^n (-1)^i \frac{1}{i!} = D_n$, lo que constituye el número de desórdenes sobre n elementos. Por consiguiente,

$$\frac{(n!)^{r-1}}{n^{n(r-2)}} D_n \prod_{i=2}^{r-1} (n - i)^n \leq L(r, n) \leq n! D_n \prod_{i=2}^{r-1} (n - i)!^{\frac{n}{n-i}}.$$

Si ponemos $r = n$, obtendremos $L(n, n) \geq (n!)^{2n} n^{-n^2} > (e^{-2}n)^{n^2}$. Se ha aprovechado aquí la conocida desigualdad: $n! > e^{-n}n^n$. Wilson [58] utilizó la desigualdad reducida $L(n, n) > (e^{-2}n)^{n^2}$ para determinar la frontera inferior del número $N(v)$ de ternas no isomorfas de Steiner de orden v y obtuvo la siguiente estimación:

$$N(v) \geq (e^{-5}v)^{v^2/6}.$$

Examinemos en conclusión las aplicaciones de los permanentes a los problemas de dímeros que surgieron en la física y química. Llamemos n -ladrillo a un paralelepípedo k -dimensional ($k \geq 2$) de volumen n , en el cual las longitudes de las aristas son de números enteros: $a_1, \dots, a_k$. Sea n par. Se plantea una cuestión: ¿cuántos son los métodos para componer un n -ladrillo, sirviéndose de 2-ladrillos (dímeros)? Denotemos este número con N . Dividamos el n -ladrillo en cubos unitarios y numeremos los: $1, 2, \dots, n$. Supongamos que $A = [a_{ij}]$ es una matriz de incidencia que caracteriza la disposición mutua de los cubos: $a_{ij} = 1$ ó 0 , lo que depende de si tienen o no aristas comunes los cubos marcados con los números i y j ($a_{11} = a_{22} = \dots = a_{nn} = 0$).

Se afirma que $N^2 = \text{per } A$. Efectivamente, a todo par ordenado de particiones del n -ladrillo en dímeros le corresponde una transversal biunívoca de la matriz A compuesta por unidades. Por ejemplo, sea $n = 6$, $k = 2$, $a_1 = 2$, $a_2 = 3$. La matriz de incidencia tendrá la forma siguiente:

1	2	3
4	5	6

Fig. 4.3.

1	2	3
4	5	6

1	2	3
4	5	6

Fig. 4.4.

	1	2	3	4	5	6
1	0	1	0	1	0	0
2	1	0	1	0	1	0
3	0	1	0	0	0	1
4	1	0	0	0	1	0
5	0	1	0	1	0	1
6	0	0	1	0	1	0

Al par ordenado de particiones del ladrillo en la fig. 4.4, le corresponde unívocamente la permutación

$$\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 4 & 1 & 6 & 5 & 2 & 3 \end{pmatrix}$$

y la transversal de la matriz A compuesta por unidades contorneadas en la fig. 4.3.

Observemos que todas las sumas respecto de las filas y columnas de la matriz de incidencia A son iguales a $2k$, a excepción de aquellas que corresponden a los cubos en la superficie. Veamos, en lugar del n -ladrillo un ladrillo toroidal correspondiente, el cual se obtiene por pegadura de las aristas opuestas del ladrillo original. La matriz de incidencia A' del ladrillo toroidal tiene $2k$ unidades en cada fila y en cada columna. Valiéndonos de la desigualdad de Mine, obtendremos: $\text{per } A' \leq \text{per } A' \leq (2k)!^{n/2k}$, y, por consiguiente, $N \leq (2k)!^{n/2k}$.

Hammersley [75] mostró que si $a_i \rightarrow \infty$ para todo $i = 1, \dots, k$, la magnitud $n^{-1} \ln N$ tiende a cierto límite λ_k . De este modo,

$$\lambda_k \leq \frac{1}{2} \ln [(2k)!]^{1/2k} < \frac{1}{2} \ln k.$$

La aplicación de la desigualdad de Van der Waerden a la matriz A' nos conduce a la estimación por debajo: $\lambda_k \geq \frac{1}{2} \ln \frac{2k}{e}$ (véase ([75])). Por consiguiente,

$$\lambda_k \sim \frac{1}{2} \ln \frac{2k}{e} (1 + o(1)) \quad \text{para } k \rightarrow \infty.$$

5.1. EVIDENCIA EN LAS INTERPRETACIONES DE LOS SISTEMAS DISCRETOS

Los sistemas discretos, a cuyos métodos de estudio se dedica este libro, existen en una diversidad exclusivamente grande. La apreciación del carácter discreto de tal o cual sistema se forma o bien debido a la comprensión de la existencia separada y discreta de los elementos del sistema dado, o bien se predetermina por el carácter discreto de la información recibida (por ejemplo, lecturas de los instrumentos). Correspondientemente, los métodos matemáticos de investigación de los sistemas semejantes se caracterizan por la misma diversidad.

Los modelos matemáticos de los sistemas discretos son, con mayor frecuencia, conjuntos de símbolos dispuestos en la forma discreta, habitualmente de puntos, unidos con líneas, si esto está determinado por las condiciones del problema.

A los modelos combinatorios se les atribuye una evidencia máxima posible no sólo para facilitar la comprensión de tal o cual problema. Los medios ilustrativos o didácticos del análisis combinatorio contribuyen frecuentemente a crear las posibilidades para resolver los problemas. De ejemplo puede servir el teorema de Ramsey (véase § 3.3).

Recordemos que el teorema de Ramsey generaliza la idea de partición de los conjuntos. Sea $P_r(S) = A_1 \cup A_2 \cup \dots \cup A_t$ una partición del conjunto de todos los r -subconjuntos del conjunto S . Supongamos, luego, que vienen dados números enteros q_i tales que $1 \leq r \leq q_i$; $i = 1, 2, \dots, t$. Si existe un q_i -subconjunto de S , cuyos r -subconjuntos se contienen todos en A_i , entonces se le llama (q_i, A_i) -subconjunto del conjunto S . El teorema de Ramsey afirma que existe un número mínimo natural $N(q_1, q_2, \dots, q_t, r)$, a partir del cual (es decir, para $n \geq N$) el conjunto S contiene un (q_i, A_i) -subconjunto para cierto $i = \{1, 2, \dots, t\}$.

Veamos un caso particular del teorema de Ramsey en que $r = 2$. La interpretación geométrica en este caso se expresa así: n elementos corresponden a n puntos en un plano; a un par de puntos corresponde un segmento que une dichos puntos; la inclusión del segmento en el subconjunto A_i se marca pintándolo de i -ésimo color ($i = 1, 2, \dots, t$). El teorema de Ramsey garantiza la existencia, para $n \geq R(p, q; 2)$, de p puntos unidos entre sí sólo con segmentos del primer color, o bien de q puntos unidos sólo con segmentos del segundo color.

La misma idea sirve de base para la aplicación del teorema de Ramsey a los polígonos convexos. Se analizan los n -conjuntos de puntos en un plano tales que ninguna combinación de tres puntos se dispone en una

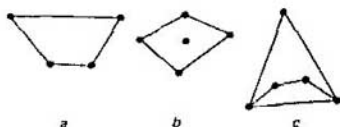


Fig. 5.1.

misma recta. Resulta que empleando los puntos de dicho conjunto como vértices, podemos construir cualquier polígono convexo de m lados, siempre que n sea suficientemente grande.

Teorema. Para un número entero dado m existe un número entero mínimo N_m tal que cualesquiera N_m puntos en el plano, de los cuales ninguna combinación de tres se dispone en una misma recta, contienen m puntos que forman un polígono convexo de m lados.

Demostración. Observemos, para iniciar, que entre cualesquiera 5 puntos se pueden elegir 4 que representen los vértices de un cuadrilátero convexo. La demostración se observa fácilmente en la fig. 5.1. Según se ve, incluso el caso "degenerado" (fig. 5.1, c) da un cuadrilátero convexo (en el interior de un triángulo).

Ahora, es lícita la afirmación de que si entre m puntos en consideración cualesquiera 4 son vértices de un cuadrilátero convexo, entonces dichos m puntos forman un polígono convexo de q lados. En efecto, supongamos que existe, a lo sumo, un polígono convexo de q lados ($q < m$). Sirviéndonos de las diagonales trazadas a partir de un mismo vértice, dividamos este polígono en triángulos y en este caso $(m - q)$ puntos caerán en el interior de los triángulos, lo que contradice la condición de convexidad de todos los cuadriláteros. Esto quiere decir que $q = m$, el polígono de m lados será convexo.

Analicemos ahora la afirmación general. Sea $m \geq 4$; $n \geq N(m, 5; 4)$. Partiremos los 4-conjuntos (hay en total $\binom{n}{4}$ de tales conjuntos) en convexos y no convexos. Entonces, según el teorema de Ramsey, o bien: a) existe un pentágono con todos los cuadriláteros no convexos, lo que contradice lo demostrado anteriormente; o bien b) existe un polígono de m lados cuyos cuadriláteros son todos convexos, y, por consiguiente, de conformidad con lo demostrado anteriormente, este polígono de m lados es convexo.

Las interpretaciones geométricas existen para casi todos los objetos combinatorios. Por ejemplo, para los cuadrados latinos la interpretación geométrica se construye del modo siguiente. Cada uno de los n^2 elementos de un cuadrado latino se considera como un punto sobre la diversidad de 2-coordenada. Por el conjunto de puntos que forman el cuadrado latino tracemos tres n -familias de líneas que pasan: 1) por las filas, 2) por las columnas, 3) por los elementos iguales. Estas familias forman una 3-red

con n^2 nodos. Inversamente, cada 3-red de esta índole puede interpretarse como cuadrado latino.

La condición de equivalencia de los cuadrados latinos en la interpretación que se considera aparece como conservación de las 3-redes al reenumerar las líneas de cada familia. A veces se admite la sustitución recíproca de las familias, con lo cual se amplía el concepto de equivalencia. Diremos, por fin, que la interpretación geométrica fue, evidentemente, la causa por la cual surgió el concepto de ortogonalidad en aplicación a los cuadrados latinos.

Una aplicación sucesiva y reiterada de las interpretaciones ilustrativas condujo a que se hayan separado clases de modelos matemáticos que poseen sus propios medios de investigación claramente distinguibles a saber, la axiomática, la terminología, los símbolos y los medios lógicos de demostración. En nuestro caso dichos medios se representan, con preferencia, por las geometrías y los grafos finitos. La tarea del capítulo presente consiste en introducir al lector en estas ramas extensas de las matemáticas, en interés del objetivo fundamental del libro.

5.2. IDEAS GEOMÉTRICAS FINITAS

Para los sistemas que se estudian en el análisis combinatorio se emplean geometrías finitas. Este término se usa para designar sistemas compuestos por un número finito de elementos, entre los cuales están establecidas las relaciones de incidencia. A los elementos, no definibles en general, se les asignan nombres geométricos: "puntos" P y "líneas" L . La relación de incidencia I se lee así: "un punto P se dispone en la línea L ", o bien "la línea L contiene el punto P ". Para dichas relaciones se introducen, además, axiomas del tipo que tienen los axiomas geométricos.

Las precisiones y explicaciones adicionales o bien las modificaciones de las enunciaciones iniciales conducen a diferentes tipos de geometrías finitas y a las partes de la geometría contiguas y próximas a las citadas geometrías. Así, por ejemplo, surgen las relaciones con los problemas de la topología combinatoria, la geometría discreta, la geometría proyectiva, la teoría geométrica de los números, la teoría de los grafos, la geometría combinatoria, etc.

Sobre los espacios proyectivos. Se conoce, por ejemplo, que la totalidad de todos los espacios de una estructura incidente $\{P, L, J\}$, que acabamos de mencionar, donde los elementos P se llaman puntos, los elementos L , líneas, y J , relaciones de incidencia, es un espacio proyectivo.

Existen diferentes métodos de construcción axiomática de un espacio proyectivo. El más empleado es la modificación del sistema de axiomas propuesta en 1899 por D. Hilbert con el fin de fundamentar la geometría elemental. Todo espacio proyectivo se considera como una colección de elementos de tres géneros: puntos, rectas y planos, entre los cuales se

establece la relación de incidencia, fundamental para la geometría proyectiva, que se caracteriza por axiomas adecuados. Estos axiomas se diferencian del grupo correspondiente de axiomas de la geometría elemental en que en los primeros se requiere que cada dos rectas dispuestas en un mismo plano tengan un punto común, y en cada recta se tengan al menos tres puntos distintos. Para obtener una geometría más sustancial, la totalidad citada de axiomas se completa con axiomas de orden y continuidad (para un espacio proyectivo real), con el axioma de Pappus (para la geometría proyectiva sobre un cuerpo conmutativo), con el postulado de Fano, etc.

Examinemos a continuación más detalladamente dos métodos de construir espacios proyectivos. En primer lugar puede tratarse de la realización de un espacio proyectivo como un conjunto de rectas en un espacio lineal. Sea L un espacio lineal sobre un campo K , y denotemos con $\dim L$ la dimensión de L . El conjunto $P(L)$ de rectas (es decir, de subespacios lineales unidimensionales) en L lleva el nombre de *espacio proyectivo* asociado con L , y las propias rectas en L se llaman puntos de $P(L)$. El número $\dim L - 1$ es la dimensión de $P(L)$ y se denota $\dim P(L)$. Los espacios proyectivos y bidimensionales se denominan *recta proyectiva* y *plano proyectivo*, respectivamente.

Un conjunto de la forma $P(A)$, donde $A \subseteq L$ se llama *subespacio proyectivo* en $P(L)$. Es evidente que $P(A_1 \cap A_2) = P(A_1) \cap P(A_2)$, es decir, una familia de espacios proyectivos está cerrada respecto de las intersecciones. Por eso, en el conjunto de subespacios proyectivos $P(A)$ que contienen el conjunto dado $S \subseteq P(L)$ hay un subespacio menor, que se representa por la intersección de todos los subespacios de este tipo. Este último lleva el nombre de *cápsula proyectiva* de S , se denota con $\overline{S}$, y coincide con $P(A)$, donde A es la cápsula lineal de todas las rectas correspondientes a los puntos $s \in S$ en L . Se puede mostrar, que si P_1 y P_2 son dos subespacios proyectivos finitos en el espacio proyectivo P , entonces

$$\dim(P_1 \cap P_2) + \dim(\overline{P_1 \cap P_2}) = \dim P_1 + \dim P_2.$$

En adelante nos harán falta dos espacios proyectivos concretos que se denominan configuraciones de Desargues y de Pappus.

Configuración de Desargues. Sea A una familia de puntos en un espacio proyectivo. El símbolo $\overline{A}$ designará su cápsula proyectiva. En un espacio proyectivo tridimensional examinemos el seis ordenado de puntos $(a_1, a_2, a_3, b_1, b_2, b_3)$. Se supone que los puntos son distintos de dos en dos y que $a_1a_2a_3$ y $b_1b_2b_3$ son, en esencia, unos planos. Luego, supongamos que las rectas a_1b_1 , a_2b_2 y a_3b_3 se intersecan en un solo punto p , distinto de $a_1, a_2, a_3, b_1, b_2, b_3$ (fig. 5.2). En este caso diremos que los triángulos $(a_1a_2a_3)$ y $(b_1b_2b_3)$ son perspectivas respecto del punto p y cada uno de ellos es una proyección del otro realizada desde el centro p , si ellos se disponen en diferentes planos. Entonces, para todo par de índices $\{i, j\} \subseteq \{1, 2, 3\}$ las rectas a_ia_j y b_ib_j no coinciden, pues en el caso contrario

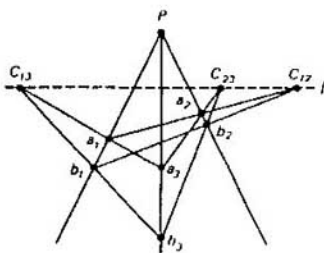


Fig. 5.2.

tendríamos $a_i = b_i$, puesto que a_i y b_i son puntos de intersección de estas rectas con la recta pa_ib_i . Además, las rectas a_ia_j y b_ib_j se disponen en el plano común pa_ia_j . Por eso, ellas se intersecan en un punto que se designará con c_{ij} . Los puntos c_{12} , c_{13} , c_{23} son puntos de intersección de las prolongaciones de los pares de los lados correspondientes de los triángulos $(a_1a_2a_3)$ y $(b_1b_2b_3)$.

Diremos que los triángulos $(a_1a_2a_3)$ y $(b_1b_2b_3)$ son perspectivos respecto de la recta l , siempre que los puntos c_{12} , c_{13} , c_{23} se disponen en la recta l .

Teorema de Desargues. Si dos triángulos son perspectivos respecto de cierto punto, son perspectivos respecto de una recta.

Configuración de Pappus. Examinemos en un plano proyectivo dos rectas diferentes l_1 , l_2 y dos ternas de puntos distintos a_1, a_2, a_3 y b_1, b_2, b_3 (fig. 5.3) dispuestos en las rectas de dos en dos. Para cualquier par de índices $\{i, j\} \in \{1, 2, 3\}$ tales que $i < j$ designemos con c_{ij} el punto de intersección de las rectas a_ib_j y a_jb_i .

Teorema de Pappus. Los puntos c_{12} , c_{13} , c_{23} se disponen en una misma recta.

El lector puede encontrar la demostración de los teoremas de Desargues y de Pappus, como también otros pormenores, en el libro [69].

En segundo lugar, los espacios proyectivos pueden construirse por adición a un espacio afín de los "puntos infinitos alejados". Merced a este procedimiento se logra la uniformidad en las enunciaciones y demostraciones de toda una serie de teoremas de la geometría analítica y de algunos

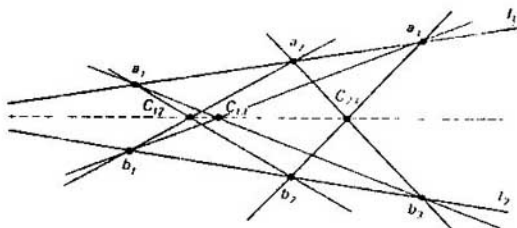


Fig. 5.3.

otros resultados geométricos. Este método condujo a la creación de la geometría proyectiva clásica. A medida que se desarrollaba el método axiomático fueron creadas diferentes descripciones axiomáticas de los espacios proyectivos. Las investigaciones ulteriores han mostrado que los sistemas geométricos que obedecen a dichos axiomas tienen su propio significado y la teoría de ellos resulta ser muy profunda (véanse [4, 21, 76 y 77]). Desde entonces por *espacio proyectivo clásico* se entiende, por lo común precisamente tal sistema geométrico. Por ejemplo, un espacio tridimensional se define como un conjunto, cuyos elementos se llaman *puntos*, provisto de dos sistemas de subconjuntos cuyos elementos se llaman *rectas* y *planos*, respectivamente. En este caso han de cumplirse los siguientes axiomas de incidencia:

1. Dos puntos distintos pertenecen a una recta única.
2. Tres puntos distintos, no dispuestos en una recta, pertenecen a un plano único.
3. Una recta y un plano tienen un punto común.
4. La intersección de dos planos contiene una recta.
5. Existen cuatro puntos que no se disponen en un plano y que son tales que cualesquiera tres de ellos no se disponen en una recta.
6. Cada recta consta no menos que de tres puntos.

En lo que sigue examinaremos sólo las geometrías proyectivas clásicas, denominándolas, para simplificar, *geometrías proyectivas*.

Un *plano proyectivo clásico* se define como un conjunto, cuyos elementos se llaman *puntos*, provisto de un sistema de subconjuntos cuyos elementos se llaman *rectas*. En este caso han de cumplirse los siguientes axiomas:

1. Dos puntos diferentes pertenecen a una recta única.
2. La intersección de dos rectas es no vacía.
3. Existen tres puntos que no se disponen en una recta.
4. Cada recta consta por lo menos de tres puntos.

El conjunto $P(L)$, donde L es un espacio lineal sobre el campo K de dimensión 4 ó 3, junto con los sistemas de planos proyectivos y rectas en ellos, tal como fueron definidos más arriba, satisfacen los axiomas del espacio proyectivo clásico tridimensional y del plano proyectivo clásico, respectivamente. No obstante, no todo espacio proyectivo clásico o plano es isomorfo a uno de nuestros espacios $P(L)$. Más aún, resulta que existen planos proyectivos clásicos no isomorfos incluso a ningún plano del tipo $P(L)$, donde L es un espacio proyectivo tridimensional sobre algún cuerpo. La razón de ello radica en que en los planos proyectivos del tipo $P(L)$ el teorema de Desargues sigue siendo válido, mientras que existen planos que no son de Desargues, donde el teorema no se cumple. Enunciamos sin demostración el resultado siguiente:

Teorema 1. Tres propiedades del plano proyectivo clásico son equivalentes:

- a) en dicho plano se cumple el teorema plano de Desargues;

b) dicho plano puede encajarse en un espacio proyectivo clásico;

c) existe un espacio lineal tridimensional L sobre cierto campo K , definido de un modo unívoco, exactamente hasta un isomorfismo tal, que nuestro plano es isomorfo a $P(L)$.

El teorema de Pappus puede fallar incluso en los planos de Desargues. Al denominar una afirmación correspondiente axioma de Pappus, podemos formular el siguiente teorema que también se da sin demostración.

Teorema 2. Si en un plano proyectivo clásico se cumple el axioma de Pappus, el plano es de Desargues. El plano clásico de Desargues satisface el axioma de Pappus, cuando y sólo cuando el cuerpo asociado con el plano es conmutativo, es decir, el plano citado es isomorfo a $P(L)$, donde L es un espacio lineal tridimensional sobre el campo.

En todo espacio proyectivo clásico de dimensión superior a dos queda válido el teorema de Desargues. El teorema de Desargues permite la coordinatización de un espacio proyectivo clásico con ayuda de cierto cuerpo asociativo K . De este modo, el estudio de los espacios proyectivos clásicos, cuya dimensión es superior a dos, se reduce a un problema algebraico, a saber, a la teoría de los cuerpos asociativos. En los planos proyectivos el teorema de Desargues puede no tener lugar y la reducción al álgebra mencionada no se consigue en el caso general. Con este motivo el estudio de los planos proyectivos clásicos es de interés especial y la teoría de ellos se desarrolla exitosamente desde el principio de nuestro siglo.

Planos proyectivos finitos. Es fácil comprobar que la definición de plano proyectivo clásico propuesta más arriba es equivalente a la siguiente.

Se denomina plano proyectivo a un conjunto de puntos y rectas, entre los cuales queda establecida una relación de incidencia que obedece a los siguientes requisitos.

1. Dos puntos distintos son incidentes, con una y sólo una recta.
2. Dos rectas distintas son incidentes al menos con un punto.
3. Existen cuatro puntos en una posición general, es decir, tales que cualesquiera tres de ellos no son incidentes con una recta.

Efectivamente, para convencerse de esta afirmación, basta demostrar el siguiente teorema.

Teorema 3. Cada recta contiene no menos de tres puntos.

Demostración. Del axioma 3 se deduce que existen cuatro puntos a_1, a_2, a_3, a_4 , de los cuales ninguna combinación de tres se dispone en una misma recta. A estos puntos les corresponden seis rectas distintas que los unen de dos en dos a saber, $l_1: a_1a_2b_1$; $l_2: a_1a_3b_2$; $l_3: a_1a_4b_3$; $l_4: a_2a_3b_4$; $l_5: a_2a_4b_5$; $l_6: a_3a_4b_6$. Aquí, b_1, b_2 y b_3 son puntos de intersección de estas rectas. Partiendo de que todas las seis rectas son distintas se obtiene con facilidad que todos los puntos b_1, b_2, b_3 son distintos y todos ellos se diferencian de los puntos a_i ; $i = 1, 2, 3, 4$. Cada una de las rectas l_i ($i = 1, 2, \dots, 6$) contiene por lo menos tres puntos. Si una recta arbitraria l no

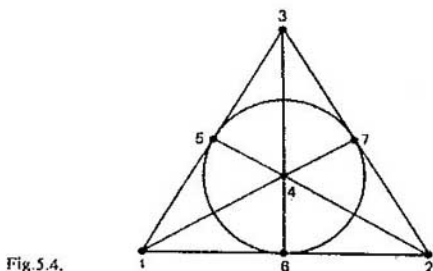


Fig. 5.4.

contiene a_1 , se interseca con las rectas l_1, l_2, l_3 en tres puntos diferentes. Si l no contiene a_2 , corta las rectas l_1, l_4, l_5 en tres puntos distintos. En cambio, si la recta l contiene tanto a_1 , como a_2 , entonces $l = l_1$, mas la recta l_1 contiene en todo caso tres puntos diferentes a_1, a_2 y b_1 . El teorema queda demostrado.

Observemos que del axioma 1 se desprende que dos rectas diferentes son incidentes no más que con un punto. En la fig. 5.4 se expone un plano proyectivo finito que contiene siete puntos y el mismo número de rectas, con la particularidad de que además de los lados y las medianas del triángulo equilátero tenemos que considerar como "recta" la circunferencia inscrita. Este plano proyectivo concreto se denomina, a menudo, configuración de Fano. Otro ejemplo de plano proyectivo ya ha sido examinado, precisamente el $P(L)$, cuando L es un espacio lineal sobre el campo de dimensión 3.

El axioma 3 en la definición de plano proyectivo sirve, principalmente, para desechar los planos "degenerados" tales, por ejemplo, como un plano vacío, una serie rectilínea de puntos (en particular, una recta), un haz de rectas (en particular, un punto), una serie rectangular de puntos por uno de los cuales están trazadas varias rectas y un haz de rectas, cortado por una recta.

Introduzcamos unas definiciones más. Un subconjunto π' de puntos y de rectas del plano proyectivo π se llama *subplano*, si, junto con cada dos puntos (rectas), π' contiene una recta (un punto) incidente con ellos. Dos planos se denominan *isomorfos*, si entre sus puntos y sus rectas puede establecerse una correspondencia biunívoca que conserva la incidencia. Específico para los planos proyectivos resulta ser el concepto de dualidad. Dos planos se llaman *duales*, si entre los puntos (rectas) de un plano y las rectas (puntos) del otro puede establecerse una correspondencia biunívoca que conserva la incidencia. Un plano dual al plano dado puede obtenerse al declarar "puntos" las rectas del plano dado, y "rectas", los puntos de él, conservando las incidencias existentes. No es difícil convencerse de que un plano isomorfo a un plano que es dual del plano π ,

será dual con relación a π . Un plano, dual respecto del dual, es isomorfo al plano de partida. La aplicación isomorfa de un plano sobre sí mismo lleva el nombre de *colineación*. La aplicación unívoca (no forzosamente biunívoca) de los puntos y de las rectas del plano π sobre los puntos y las rectas del plano π' con la conservación de la incidencia recibe el nombre de homomorfismo.

Teorema 4. Sea $n \geq 2$ un número entero arbitrario. Entonces son equivalentes las siguientes propiedades del plano proyectivo:

- a) cierta recta contiene exactamente $n + 1$ puntos;
- b) cierto punto pertenece exactamente a $n + 1$ rectas;
- c) cada recta contiene exactamente $n + 1$ puntos;
- d) cada punto se dispone exactamente en $n + 1$ rectas;
- e) en el plano π hay exactamente $n^2 + n + 1$ puntos;
- f) en el plano π hay exactamente $n^2 + n + 1$ rectas.

Demostración. Sean a_1, a_2, a_3, a_4 cuatro puntos, de los cuales ninguna combinación de tres se dispone en una misma recta. A estos puntos les corresponden seis diferentes rectas que los unen dos a dos (véase la demostración del teorema 3). Supongamos ahora que se cumple la propiedad a), es decir, en π existe una recta l que contiene exactamente $n + 1$ puntos, digamos, para concretar, $c_1, c_2, \dots, c_{n+1}$. Si b es un punto que no está dispuesto en l , entonces las rectas $bc_i, i = 1, 2, \dots, n + 1$, son distintas, pues, si $\overline{bc_i} = \overline{bc_j}$ para ciertos $i \neq j$, entonces b está situado en l , lo que contradice la hipótesis. Luego, cada recta, que pasa por b , corta l , y, por consiguiente, ha de ser una de las $n + 1$ rectas $bc_i, i = 1, 2, \dots, n + 1$. Por lo menos dos puntos de los cuatro a_1, a_2, a_3, a_4 no se disponen en la recta l , y, por eso, tal punto b existe. Sea, ahora, b un punto que yace exactamente en $n + 1$ rectas: $m_1, m_2, \dots, m_{n+1}$. Si l^* es una recta que no pasa por b , entonces l^* corta $m_1, \dots, m_{n+1}$ en los puntos $d_1, \dots, d_{n+1}$, que son todos diferentes, por cuanto b es el único punto que yace más que en una de las rectas $m_1, \dots, m_{n+1}$. Si en l^* existiera un punto más d_{n+2} , entonces existiría una recta bd_{n+2} que no coincide con ninguna recta $m_j, j = 1, 2, \dots, n + 1$, puesto que de lo contrario la recta bd_{n+2} contendría cierto punto $d_j, j \neq n + 2$, y, entonces, $\overline{bd_{n+2}} = \overline{bd_{n+2}d_j} = \overline{d_{n+2}d_j} = l^*$, lo que contradice la suposición de que b no está situado en la recta l^* .

Nuestra recta inicial l contenía exactamente $n + 1$ puntos; por consiguiente, cada punto que no yace en l , está situado exactamente en $n + 1$ rectas; a los puntos de esta índole se relacionan por lo menos dos puntos de cuatro: a_1, a_2, a_3, a_4 , por ejemplo, a_1 y a_2 . Por eso, cada recta que no pasa por a_1 ó a_2 contiene exactamente $n + 1$ puntos, es decir, cada recta, salvo, quizás, la recta $l_1: a_1a_2b_1$ contiene exactamente $n + 1$ puntos. Entonces, $l_2: a_1a_3b_2$ contiene exactamente $n + 1$ puntos, y el punto b_3 que no yace en l_2 , está situado exactamente en $n + 1$ rectas; por consiguiente,

l_1 , que no contiene b_3 , también debe contener $n + 1$ puntos. Así pues, la propiedad a) lleva consigo la propiedad c).

Pero, para cualquier punto b puede encontrarse una recta que no pasa por este punto, y por esta razón existen (igual que anteriormente) exactamente $n + 1$ rectas que pasan por b , con lo que quedan demostradas las propiedades b) y d).

Demostremos ahora la propiedad e), partiendo de a). Sea b_0 un cierto punto del plano π , y supongamos que $l_1, l_2, \dots, l_{n+1}$ representan $n + 1$ rectas que pasan por b_0 . Estas rectas contienen todos los puntos del plano π , con la particularidad de que cada una de ellas contiene otros b_0 y n puntos. El punto b_0 es el único, perteneciente a cualesquiera dos de las rectas $l_1, l_2, \dots, l_{n+1}$. Por consiguiente, el plano π contiene $1 + (n + 1)n = n^2 + n + 1$ puntos. Confirmemos ahora la propiedad f). Sea l_0 una recta del plano π , y supongamos que $b_1, \dots, b_{n+1}$ son $n + 1$ puntos de la recta l_0 . Cada uno de estos puntos yace en l_0 y en otras n rectas. De este modo, obtenemos todas las rectas del plano π ; en total tenemos, por consiguiente, $1 + (n + 1)n = n^2 + n + 1$ puntos. Así pues, de la propiedad a) se deducen todas las demás propiedades mencionadas en el teorema.

En virtud de la dualidad, de b) también se deducen todas las propiedades restantes. Es evidente que de c) se deduce a), y de d) se deduce b). Si es cierta la propiedad e) y cierta recta consta de $m + 1$ puntos, donde m es un número natural, entonces el plano π se compone de $m^2 + m + 1 = n^2 + n + 1$ puntos, de donde $m = n$, es decir, de e) se desprende a). Análogamente, de f) se deduce la propiedad b). El teorema está demostrado.

Convengamos en llamar *orden* de un plano proyectivo finito al número n , si cada recta de este plano contiene $n + 1$ puntos. Observemos que el plano es de orden n , si posee cualquiera de las seis propiedades citadas en el teorema 4. El orden mínimo de los planos proyectivos finitos es igual a 2. El plano finito "más pequeño" es la configuración de Fano, cada recta de la cual contiene exactamente 3 puntos. La configuración se expone en la fig. 5.4. En total dicho plano contiene 7 ($= 2^2 + 2 + 1$) puntos y el mismo número de rectas. Las rectas en este plano son: l_1 : 1, 2, 6; l_2 : 1, 3, 5; l_3 : 2, 3, 7; l_4 : 1, 4, 7; l_5 : 3, 4, 6; l_6 : 2, 4, 5; l_7 : 5, 6, 7.

Ya en este ejemplo más sencillo se revelan las posibilidades de interpretar los planos proyectivos finitos. A saber, un plano proyectivo finito de orden 2 resulta ser un sistema de ternas de Steiner de orden 7. Hemos de notar que este sistema es un bloque-esquema con los parámetros $v = b = 7$, $k = r = 3$, $\lambda = 1$. En el caso general los planos de orden n representan bloque-esquemas, cuyos parámetros tienen la forma: $v = b = n^2 + n + 1$, $k = r = n + 1$, $\lambda = 1$, es decir, bloque-esquemas incompletos equilibrados simétricos. Viceversa, los bloque-esquemas con tales parámetros son planos

proyectivos simétricos finitos, puesto que, evidentemente, se cumplen todos los axiomas. Establezcamos una conexión útil más entre los planos proyectivos finitos y las familias completas de cuadrados latinos ortogonales.

Teorema 5. Sea $n \geq 3$ un número entero arbitrario. Un plano proyectivo de orden n existe, cuando y sólo cuando puede construirse una familia completa de $n - 1$ cuadrados latinos ortogonales de orden n .

Demostración. Sea dado un plano proyectivo finito π de orden n . Veamos en él una recta l . En esta recta se tienen, como sabemos, $n + 1$ puntos: $a_1, a_2, \dots, a_{n+1}$. Supongamos que $b_1, b_2, \dots, b_n$ son los demás n^2 puntos del plano π que no están situados en l . Cada punto a_j ($j = 1, 2, \dots, n + 1$) yace en n rectas, sin tener en cuenta la recta l . Numeremos arbitrariamente con números naturales de 1 hasta n todas las n rectas que pasan por el punto a_j . Procedamos de este modo para todo $j = 1, 2, \dots, n + 1$. Supongamos que la recta $b_i a_j$ está numerada con un número natural a_{ij} . Entonces, $A = [a_{ij}]$, donde $i = 1, 2, \dots, n^2$; $j = 1, 2, \dots, n + 1$, es una $n^2 \times (n + 1)$ -tabla de elementos $1, 2, \dots, n$.

Las filas de cualquier $(n^2 \times 2)$ -subtabla elegida de A representan n^2 pares de los elementos $1, 2, \dots, n$. Si suponemos que $a_{ij} = a_{i'j}$ y $a_{ik} = a_{i'k}$, donde $i \neq i'$; $j \neq k$, entonces $b_i a_j = b_{i'} a_j$, y $b_i a_k = b_{i'} a_k$. Mas, en este caso la recta $b_i b_{i'}$ contiene tanto el punto a_j , como el a_k , es decir, la recta $b_i b_{i'}$ coincide con la recta l , lo que contradice nuestra suposición. Esto quiere decir que la matriz $A = [a_{ij}]$ es una $n^2 \times (n + 1)$ -tabla ortogonal de fuerza 2, la cual, según lo demostrado en el capítulo 4, es equivalente al conjunto completo de $n - 1$ cuadrados latinos ortogonales de orden n .

Demostremos la afirmación inversa. Sea dado un conjunto completo de $n - 1$ cuadrados latinos ortogonales de orden n . Pasemos de éste, como lo hicimos en el cap. 4, a una tabla

$$A = [a_{ij}]; \quad i = 1, 2, \dots, n^2; \quad j = 1, 2, \dots, n + 1.$$

A cada una de n^2 filas de la tabla A pongamos en correspondencia los puntos $b_1, b_2, \dots, b_{n^2}$, y a las columnas, los puntos $a_1, a_2, \dots, a_{n+1}$. Definamos la recta l_i como compuesta por los puntos $a_1, a_2, \dots, a_{n+1}$. Por todo punto a_j tracemos n rectas l_{ij} que contienen aquellos puntos b_i , donde en la j -ésima columna figura el número i . Se obtiene un plano proyectivo finito de orden n , en el cual se tienen $n^2 + n + 1$ rectas y el mismo número de puntos, con la particularidad de que cada recta contiene exactamente $n + 1$ puntos y cada punto yace en $n + 1$ rectas. Efectivamente, sean l_{ij} y $l_{i'j}$ dos rectas, y, además $i \neq i'$. Estas rectas son incidentes con uno y sólo con un punto a_j . Las rectas l_{ij} y l también son incidentes con uno y sólo con un punto. Esto demuestra el requisito 3 en la definición de plano proyectivo. Los cuatro puntos: $(1, 1), (1, 2), (2, 1), (2, 2)$ satisfacen el requisito 3 de la definición de plano proyectivo, lo que quiere decir que

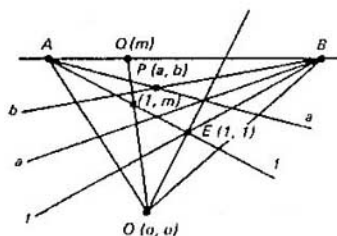


Fig. 5.5.

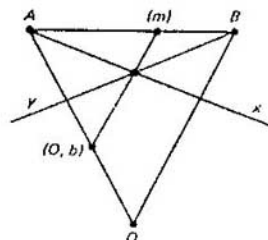


Fig. 5.6.

el plano construido π es realmente un plano proyectivo de orden n . El teorema queda completamente demostrado.

Introducción de las coordenadas. Elijamos en un plano proyectivo π cuatro puntos A, B, O, E en la posición general (véase fig. 5.5). Llamemos haz A a un conjunto de rectas que pasan por el punto a . Pongamos a cada recta del haz A en correspondencia aquella recta del haz B con la que ella se encuentra en la recta OE . Atribuyamos a cada una de las rectas del haz A (a excepción de AB) un símbolo especial a y asignemos el mismo símbolo a la recta correspondiente del haz B . Agreguemos a la recta AO símbolo 0 y a la recta AE , el símbolo 1 . Las rectas BO y BE recibirán los mismos símbolos. De este modo se define cierto conjunto de símbolos $M = \{0, 1, \{a\}\}$.

Supongamos que el punto P no pertenece a la recta AB (fig. 5.5). Por dicho punto pasa una recta de cada haz A y B , a saber, las rectas AP y BP . Supongamos que a la recta AP se le asigna el símbolo a , y a BP , el símbolo b . Entonces, al punto P le ponemos en correspondencia un par ordenado de símbolos (a, b) de M , que se denominarán coordenadas de este punto. Es fácil ver que la correspondencia entre los puntos que no están situados en AB y los pares (a, b) será biunívoca. Además, los puntos dispuestos en OE tendrán iguales ambas coordenadas. Asignemos ahora las coordenadas a los puntos de la recta AB distintos de A . Con este fin unamos el punto Q (véase fig. 5.5) con O y veamos donde la recta QO se cortará con AE . La segunda coordenada del punto de intersección se tomará por la coordenada del punto Q .

Observemos que mediante la misma construcción pueden obtenerse coordenadas cartesianas en un plano real habitual. El punto A debe considerarse como punto impropio del eje y , el punto B , como punto impropio del eje x , O se toma por origen de coordenadas y E , como el punto $(1, 1)$.

Definamos sobre el conjunto M , obtenido como resultado de la coordenatización, una operación ternaria, es decir, a cada terna ordenada de elementos x, m, b de M le ponemos en correspondencia un cierto elemento

$$y = x \cdot m \circ b$$

dispuesto en M . Este elemento y se define a base de los elementos x, m, b , como segunda coordenada del punto de intersección de la recta x del haz A con la recta (m) , $(0, b)$ (fig. 5.6). Es evidente que por medio de este procedimiento y se define de un modo unívoco.

Teorema 6. Cualquier forma de prefijar cuatro puntos A, B, O, E en una posición general determina una operación ternaria con las siguientes propiedades:

1. $0 \cdot m \circ c = a \cdot 0 \circ c = 0$.
2. $1 \cdot m \circ 0 = m \cdot 1 \circ 0 = m$.
3. $a \cdot m \circ z = C$ resoluble unívocamente respecto de z .
4. $x \cdot m_1 \circ b_1 = x \cdot m_2 \circ b_2$ resoluble unívocamente respecto de x .
5. El sistema

$$a_1 = m \circ b = c_1; a_2 \cdot m \circ b = c_2; a_1 = a_2$$

define unívocamente un par (m, b) .

Demostración. Al elegir cuatro puntos A, B, O, E , de los cuales ninguna combinación de tres yace en una recta del plano π , formamos un anillo ternario con la operación $x \cdot m \circ b$, al igual que lo hicimos más arriba. Las propiedades 1 y 2 se deducen inmediatamente de la definición. La propiedad 3 significa que una recta que liga los puntos (m) y (a, c) corta OB en un punto bien determinado $(0, z)$. El significado de la exigencia 4 consiste en que dos rectas: $y = x \cdot m_1 \circ b_1$ e $y = x \cdot m_2 \circ b_2$ con diferentes direcciones m_1 y m_2 se intersecan en el único punto finito. La exigencia 5 dice que si (a_1, c_1) y (a_2, c_2) son dos puntos finitos para $a_1 \neq a_2$, entonces existe una recta única $y = x \cdot m \circ b$, que pasa por estos dos puntos.

El conjunto M con la operación ternaria que posee las propiedades 1...5 del teorema 6 lleva el nombre de *terna*. Si una terna dada está construida sobre el plano π según la regla descrita anteriormente, se llamará terna del plano π .

Aduzcamos el resultado siguiente sin demostrarlo.

Teorema 7. Dada una terna M , puede determinarse con su ayuda un plano proyectivo con los puntos (a, c) , (m) , A , y con las rectas $y = x \cdot m \circ b$, $x = a, l_\infty$, donde a, b, c, m representan toda clase de elementos de M , mientras que A y l_∞ son ciertos símbolos auxiliares; la incidencia se define del modo siguiente: (a, c) es incidente con $y = x \cdot m \circ b$, cuando y sólo cuando $c = a \cdot m \circ b$; (a, c) es incidente con $x = a$ para todo c ; (m) es incidente con $y = x \cdot m \circ b$ para todo b ; (m) es incidente con l_∞ , cualquiera que sea m ; A es incidente con $x = a$ para cualesquiera a , y, además, con l_∞ .

Introduzcamos, para el plano dado π , las operaciones naturales: adición y multiplicación, que se definen partiendo de una de las ternas de este plano, según las reglas siguientes:

$$\begin{aligned} a + b &= a \cdot 1 \circ b; \\ ab &= a \cdot b \circ 0. \end{aligned}$$

El conjunto M con la adición y multiplicación definidas del modo citado recibe el nombre de *cuerpo natural* del plano π .

Ha de notarse que un cuerpo natural no define, hablando en general, la terna que lo engendra, y el teorema, análogo al teorema 7, no tiene lugar en este caso.

Así pues, los planos proyectivos finitos admiten diferentes interpretaciones: mediante ternas, bloque-esquemas y familias completas de cuadrados latinos ortogonales. Sin embargo, la teoría de los planos finitos todavía está lejos de ser elaborada. Ni siquiera existe la respuesta completa a la pregunta sobre el volumen de los objetos combinatorios de esta clase, como también sobre las condiciones, bajo las cuales los planos finitos pueden existir y no existir.

Se conoce que un plano proyectivo finito existe, si su orden n tiene la forma p^α , donde p es un número primo, y α , un número natural. En efecto, se ha demostrado (véase cap. 4) que si $n = p^\alpha$, $n \geq 3$, entonces existe una familia completa de $n - 1$ cuadrados latinos ortogonales. Mientras tanto, la existencia de tal conjunto es equivalente a la existencia del plano proyectivo finito que se busca.

De los demás resultados señalaremos los siguientes: a) el plano finito no existe, si $n \equiv 1, 2 \pmod{4}$ y si la parte de n , libre del cuadrado, tiene por lo menos un factor $p \equiv 3 \pmod{4}$; b) la condición necesaria de existencia de un plano proyectivo finito cuyo orden tiene la forma $n \equiv 1, 2 \pmod{4}$ consiste en la existencia de los números enteros x y y tales, que $n = x^2 + y^2$ (teorema de Bruck—Ryser). No damos aquí las demostraciones por ser éstas demasiado complejas. De los resultados citados se desprende, en particular, que los planos para $n = 6, 12, 14$ están ausentes. La cuestión referente a $n = 10, 18$ y 20 , queda por ahora abierta. Para que el lector perciba lo difícil del problema, indiquemos, como ejemplo, que la construcción de un plano de orden 10 requiere la construcción de una familia de 9 cuadrados latinos ortogonales dos a dos del mismo orden.

Una dirección importante en el desarrollo de la teoría de los planos proyectivos finitos la representan el cálculo del número de planos de un orden dado y el estudio de sus tipos diferentes. Para el año 1956 quedó demostrado que cuando $n < 9$ ($n = 2, 3, 4, 5, 7, 8$) existe el único plano proyectivo de orden n , a saber, el de Desargues. Tres planos construidos por Veblen y Wedderburn en 1907, junto con el plano de Desargues construido un año antes por Veblen y Bassi son los únicos que hoy día se conocen para el orden 9. No se han obtenido planos nuevos, como tampoco se ha demostrado que ellos no existen en general. En esta dirección se emprenden varias investigaciones. Parece que estamos a punto de terminar la resolución de la cuestión sobre la existencia de los planos proyectivos de orden 10.

5.3. SOBRE LOS GRAFOS

En el lenguaje de la teoría de los grafos se describen con éxito muchos tipos de los problemas combinatorios. Las representaciones gráficas son en este caso no simplemente ilustraciones, sino que también permiten obtener nuevos resultados. En este párrafo daremos una información indispensable de la teoría de los grafos que se emplea en la obra dada.

Se denomina *grafo orientado* una terna $G = (X, A, \varphi)$ compuesta de un conjunto no vacío X , cuyos elementos llevan el nombre de *vértices* de un conjunto A de *arcos* (llamados también), a veces, flechas y de una función $\varphi: A \rightarrow X \times X$, la cual a todo arco $a \in A$ se le hace corresponder un par ordenado (p, q) de vértices denominados *inales* de dicho arco. Un arco, cuyos finales (p, p) se encuentran en un mismo vértice, lleva el nombre de *lazo*. Un grafo privado de tales lazos se llama *grafo sin lazo*.

En la fig. 5.7. están expuestos cuatro grafos orientados; cada uno con cuatro vértices. Todos estos grafos no tienen lazos. Además, son *simples*: lo último significa por definición que cualquier par de vértices p, q se une mediante un arco a lo sumo. En los grafos orientados simples el único arco (si existe) con las finales, p, q se designará mediante pq .

Describamos dos familias importantes de los grafos orientados simples, cada uno de los cuales depende del parámetro n . Un camino *simple* de longitud n consta de $n + 1$ vértices: $x_1, \dots, x_{n+1}$ y n arcos que unen los vértices vecinos: $x_k x_{k+1}$. El *contorno simple* de longitud n consta de n diferentes vértices: $x_1, \dots, x_n$ y n arcos del tipo $x_k x_{k+1}$ para $k < n$, y, además, $x_n x_1$ (en particular, un contorno simple de longitud 1 es un lazo con vértice x_1). Todos los caminos de longitud n son isomorfos, como lo son también todos los contornos simples de longitud n (véase la definición más abajo). En la fig. 5.7, *a* está expuesto un trayecto simple de longitud 3, y en la fig. 5.7, *b*, un ciclo simple de longitud 4.

Se denomina *isomorfismo* de los grafos orientados $G = (X, A, \varphi)$ y $G^* = (X^*, A^*, \varphi^*)$ a un par de biyecciones $\beta: X \rightarrow X^*$ y $j: A \rightarrow A^*$ tal que en G el arco a va del vértice p al vértice q , cuando y sólo cuando en G^* el arco $j(a)$ va del vértice $\beta(p)$ al vértice $\beta(q)$. En otras palabras, $\varphi(a) = (p, q)$ en G es equivalente a la condición $\varphi^*(j(a)) = (\beta(p), \beta(q))$ en G^* . Dos grafos orientados se llaman isomorfos, si entre ellos existe un isomorfismo. Es natural identificar los grafos orientados isomorfos.

Todo grafo orientado $G = (X, A, \varphi)$ define en el conjunto de sus vértices una relación binaria de *sucesiones* σ . Por definición, $p\sigma q$ significa que $\varphi(a) = (p, q)$ para cierto $a \in A$. Viceversa, toda relación binaria ρ en el conjunto X define un grafo orientado simple $G(\rho) = (X, A(\rho), \psi)$, en el que $A(\rho) = \{(p, q) | p \in X, q \in X, p\rho q\}$ y $\psi(pq) = (p, q)$. De aquí se deduce inmediatamente que la clase de conjuntos con una relación binaria (x, ρ) se encuentra en una correspondencia biyectiva natural con la clase de grafos orientados simples.

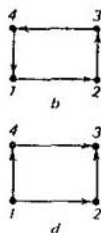
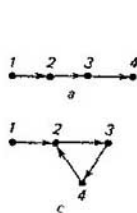


Fig. 5.7.

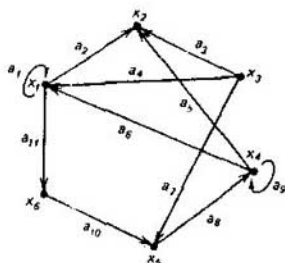


Fig. 5.8.

Se llama *grafo no orientado* (o, simplemente, *grafo*) la terna $G = (X, E, \Theta)$ compuesta de un conjunto no vacío X (cuyos elementos llevan el nombre de *vértices*), el conjunto E de aristas y la función Θ que a cada arista $a \in E$ le pone en correspondencia un par no ordenado de vértices $(p, q) = (q, p)$ que se denominan extremos de dicha arista. La arista (p, q) se llama *lazo*.

Los conceptos, introducidos para los grafos orientados, pueden ser extendidos a los grafos no orientados, si consideramos que una arista no orientada (p, q) corresponde a un par de arcos pq y qp .

Una *cadena simple* de longitud n consta de $n + 1$ diferentes vértices $x_1, \dots, x_{n+1}$ y de n aristas que unen los vértices vecinos. Un *ciclo simple* de longitud n consta de n diferentes vértices: $x_1, \dots, x_n$ y n aristas del tipo (x_k, x_{k+1}) para $k < n$, y, además, de una arista (x_n, x_1) .

El grafo orientado G puede también considerarse como un par $G = (X, \Gamma)$, formado por el conjunto X y la aplicación multiforme del conjunto X en sí mismo (es decir, por la aplicación de X en el conjunto $P(X)$ de todos los subconjuntos del conjunto X). Para el grafo expuesto en la fig. 5.8 tenemos: $\Gamma(x_3) = \{x_2, x_5\}$; $\Gamma(x_2) = \emptyset$. Con $\Gamma^{-1}(x_k)$ se designará el conjunto de vértices x_k del grafo, para los cuales en G existe un arco (x_k, x_i) . Sea $X_q = \{x_1, \dots, x_q\}$. Entonces, por $\Gamma(X_q)$ se entiende la unión $\Gamma(x_1) \cup \dots \cup \Gamma(x_q)$. La aplicación $\Gamma(\Gamma(x_i))$ se escribe como $\Gamma^2(x_i)$. Análogamente, una aplicación "triple" $\Gamma(\Gamma(\Gamma(x_i)))$ se escribe como $\Gamma^3(x_i)$, etc. Por ejemplo, para nuestro grafo orientado tenemos

$$\Gamma^2(x_1) = \Gamma(\Gamma(x_1)) = \Gamma(\{x_1, x_2, x_6\}) = \{x_1, x_2, x_5, x_6\};$$

$$\Gamma^3(x_1) = \Gamma(\{x_1, x_2, x_5, x_6\}) = \{x_1, x_2, x_1, x_5, x_6\}.$$

Un grafo (X', A') se llama *subgrafo* del grafo (X, A) , si X' y A' están contenidos en X y A , respectivamente. Se denomina *subgrafo de esqueleto* (*grafo parcial*) G_p del grafo $G = (X, A)$ al grafo (X, A_p) , para el cual $A_p \subseteq A$. De este modo, el subgrafo de esqueleto tiene también el mismo conjunto de vértices que el grafo G , pero el conjunto de arcos del

subgrafo G_p es un subconjunto del conjunto de arcos del grafo inicial. Sea un grafo $G = (X, \Gamma)$. Se llama *subgrafo generado* G_s al grafo (X_s, Γ_s) , para el cual $X_s \subseteq X$ y $\Gamma_s(x_i) = \Gamma(x_i) \cap X_s$ para cada vértice $x_i \in X_s$. Así pues, un subgrafo generado consta de un subconjunto de vértices X_s , de un conjunto de vértices del grafo inicial y de todos aquellos arcos del grafo G , cuyas vértices extremales y originales pertenecen al subconjunto X_s . En la fig. 5.9, *a, b* se exponen un subgrafo generado, un subgrafo de esqueleto y un subgrafo, respectivamente, con la particularidad de que este último no es ni generado ni subgrafo de esqueleto del grafo G , expresado en la fig. 5.8.

Sea $G = (X, A)$ un grafo orientado con un conjunto de vértices X y un conjunto de arcos A . Si queremos menospreciar el carácter orientado de los arcos del conjunto A , entonces un grafo no orientado, correspondiente a G , se denotará con $\bar{G} = (X, A)$ y se llamará *sosia (doble) no orientado* del grafo G . Un grafo no orientado G se denomina *conexo*, si cada par de sus vértices puede ser unido con una cadena. Un grafo finito, que no es conexo, puede partirse en un número finito de subgrafos finitos llamados *componentes*. Un grafo no orientado G se llama *completo*, si para cada par de sus vértices existe una arista que los une. Un grafo completo sobre n vértices suele designarse con K_n . Un grafo no orientado se denomina *bipartido*, si el conjunto de sus vértices puede ser partido en tales dos subconjuntos X_1 y X_2 que cada arista tenga uno de sus extremos en X_1 , y el otro en X_2 ; aquí $X_1 \cap X_2 = \emptyset$. Un grafo bipartido $G = (X_1 \cup X_2, A)$ se llama *completo*, si para cualesquiera dos vértices $x_i \in X_1$ y $x_j \in X_2$ existe una arista $(x_i, x_j) \in A$. Si $|X_1| = n$, y $|X_2| = m$, el grafo completo no orientado $G = (X_1 \cup X_2, A)$ se denota con $K_{n,m}$. Un grafo orientado G se llama *conexo (bipartido, respectivamente)*, si lo es su sosia no orientada $\bar{G}$.

Se llama *grado* del vértice en un grafo no orientado al número de aquellas aristas de cuyo extremo sirve el vértice.

Examinemos un grafo no orientado $G = (X, A)$ sin lazos. Elijamos una arista $a = (x, y) \in A$. Sustituyamos el vértice y en todas las aristas $b \in (A \setminus \{a\})$ por el vértice x y denotemos con A' el conjunto obtenido de aristas. Diremos que el grafo $G' = (X \setminus \{y\}, A')$ se ha obtenido del grafo G como resultado de contraer la arista a al vértice x . Un grafo G se denomina *planario*, si puede ser dibujado en un plano de un modo tal que dos aristas arbitrarias del grafo no se intersequen. En la fig. 5.10 están expuestos el grafo completo K_5 y el grafo completo bipartido $K_{3,3}$, los cuales, como se sabe, no son planarios. Estos dos grafos tienen gran importancia en la teoría de los grafos planarios y se conocen como grafos de Kuratovski, por cuanto fue él el primero en establecer que un grafo no orientado G no es planario, si y sólo si contiene un subgrafo contraído al grafo K_5 ó al $K_{3,3}$.

Matriz de adyacencia. Sea dado un grafo orientado simple G con los vértices $x_1, x_2, \dots, x_n$. Se llama *matriz de adyacencia* de dicho grafo a

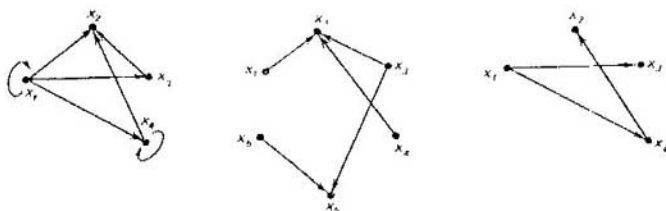


Fig. 5.9.

una matriz cuadrada $B = [b_{ij}]$ de orden n , donde

$$b_{ij} = \begin{cases} 1, & \text{si en } G \text{ existe el arco } (x_i, x_j); \\ 0, & \text{si en } G \text{ no hay arco } (x_i, x_j). \end{cases}$$

Así pues, la matriz de adyacencia del grafo expuesto en la fig. 5.8, tiene la forma siguiente

$$B = \begin{matrix} & \begin{matrix} x_1 & x_2 & x_3 & x_4 & x_5 & x_6 \end{matrix} \\ \begin{matrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{matrix} & \begin{pmatrix} 1 & 1 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix} \end{pmatrix}$$

La matriz de adyacencia define completamente la estructura del grafo. Por ejemplo, la suma de todos los elementos de la fila x_i de la matriz B da el número de arcos que tienen el vértice x_i como su vértice original, y la suma de elementos de la columna x_i da el número de arcos en los que x_i figura como vértice final. El conjunto de columnas que tienen 1 en la fila x_i es un conjunto $\Gamma(x_i)$, y el conjunto de filas que tienen 1 en la columna x_i coincide con el conjunto $\Gamma^{-1}(x_i)$.

Elevemos al cuadrado la matriz de adyacencia. Sea $b_{ik}^{(2)}$ un elemento de la matriz B^2 , entonces, $b_{ik}^{(2)} = \sum_{j=1}^n b_{ij}b_{jk}$. El sumando $b_{ij}b_{jk}$ es igual a 1, cuando y sólo cuando ambos números b_{ij} y b_{jk} son iguales a 1; de lo contrario, es igual a 0. De la igualdad $b_{ij} = b_{jk}$ se desprende la existencia del

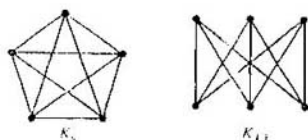


Fig. 5.10

camino de longitud 2 del vértice x_i al vértice x_k , pasando por el vértice x_j . Por consiguiente, $b_{ik}^{(2)}$ en la matriz B^2 es igual al número de caminos de longitud 2 que van de x_i a x_k .

Análogamente, el elemento $b_{ik}^{(p)}$ de la matriz B^p es igual al número de caminos (no forzosamente simples) de longitud p que van de x_i a x_k .

La matriz de adyacencia de un grafo no orientado se construye de un modo análogo.

Matriz de incidencias. Sea dado un grafo orientado G con los vértices $x_1, x_2, \dots, x_n$ y los arcos $a_1, a_2, \dots, a_m$. Se llama matriz de incidencias del grafo G a una matriz $A = [a_{ij}]$ de dimensión $n \times m$, cuyas filas corresponden a los vértices, y las columnas, a los arcos, y

$$a_{ij} = \begin{cases} 1 & \text{si el vértice } x_i \text{ es el origen del arco } a_j; \\ -1, & \text{si el vértice } x_i \text{ es el extremo del arco } a_j; \\ 0, & \text{si el vértice } x_i \text{ no es incidente con el arco } a_j, \text{ o si el arco } a_j \text{ no es un lazo.} \end{cases}$$

Por ejemplo, para el grafo que se expone en la fig. 5.8 la matriz de incidencias es de la forma

$$A = \begin{matrix} & \begin{matrix} a_1 & a_2 & a_3 & a_4 & a_5 & a_6 & a_7 & a_8 & a_9 & a_{10} & a_{11} \end{matrix} \\ \begin{matrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{matrix} & \left\{ \begin{array}{cccccccccccc} 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & -1 & -1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 \end{array} \right\} \end{matrix}$$

Por cuanto cada arco es incidente con dos diferentes vértices, a excepción del caso en que un arco forma un lazo, cada columna contiene o bien un solo elemento igual a 1, y un elemento igual a -1, o bien todos los elementos iguales a 0.

Si G es un grafo no orientado, su matriz de incidencias se determina igual que antes, salvo que todos los elementos iguales a -1 se sustituyen por +1.

Matriz de ciclos (de contornos). Sea G un grafo no orientado. Demos a sus aristas los números $a_1, a_2, \dots, a_m$, y a los ciclos simples, los números $c_1, c_2, \dots, c_p$. Se llama matriz de ciclos del grafo G a una matriz $C = [c_{ij}]$ de dimensión $p \times m$, donde

$$c_{ij} = \begin{cases} 1, & \text{si el ciclo simple } c_i \text{ del grafo } G \text{ contiene la arista } a_j; \\ 0, & \text{en el caso contrario.} \end{cases}$$

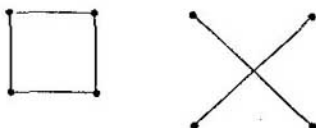


Fig. 5.11.

Así pues, las filas de una matriz de ciclos corresponden a los ciclos simples, y las columnas, a las aristas.

Si G es un grafo orientado, numeremos sus arcos con $a_1, a_2, \dots, a_m$, y los contornos simples, con $c_1, c_2, \dots, c_p$; la matriz de contornos $C = [c_{ij}]$ se determina de un modo análogo:

$$c_{ij} = \begin{cases} 1, & \text{si el contorno } c_i \text{ contiene un arco } a_j \text{ y los mismos est\'an igualmente orientados;} \\ -1, & \text{si el contorno } c_i \text{ contiene un arco inverso de } a_j; \\ 0, & \text{en todos los dem\'as casos.} \end{cases}$$

Observaci\'on. La elecci\'on de s\'ımbolos iguales para denotar las matrices de los grafos orientado y no orientado se debe a que tanto los grafos orientados, como los no orientados, poseen iguales propiedades en lo que se refiere a la estructura de estas matrices. Por eso, no hay confusi\'on ninguna al introducir tales s\'ımbolos. Adem\'as, las propiedades de las matrices de los grafos orientados y los m\'etodos de su demostraci\'on son los mismos que para las matrices de los grafos no orientados.

Ejercicios. Cercior\'ese de que son v\'alidas las siguientes afirmaciones.

1. En todo grafo no orientado finito el n\'umero de v\'ertices de grado impar es par.
2. Sea A una matriz de incidencias de un grafo no orientado simple sin lazos. Entonces, la matriz de adyacencia B se obtiene de $A \cdot A^T$ (donde A^T es una matriz obtenida por transposici\'on de la matriz A) por sustituci\'on de todos los elementos en la diagonal por ceros.
3. Un grafo G es bipartido, cuando y s\'olo cuando para todo n\'umero impar n todos los elementos diagonales de la matriz A^n son nulos.
4. Si los grafos G_1 y G_2 son isomorfos, los valores propios de las matrices de adyacencia de estos grafos son iguales.

Observaci\'on. La afirmaci\'on rec\'ıproca no es v\'alida en el caso general. Por ejemplo, dos grafos expuestos en la fig. 5.11 tienen iguales valores propios ($\lambda = \pm 2, 0, 0, 0$), sin embargo, evidentemente, ellos no son isomorfos.

5. Cualquier menor de la matriz de incidencias del grafo G es igual a $\pm 1, -1$, o a 0 .
6. El rango de una matriz de incidencias A de un grafo conexo (tanto orientado, como no orientado) G con n v\'ertices es igual a $n - 1$.
7. El rango de una matriz de ciclos C de un grafo conexo G con m aristas y n v\'ertices es igual a $m - n + 1$.
8. Una matriz de incidencias A es ortogonal respecto de la matriz C^T , transpuesta a la matriz de ciclos C y obtenida con el mismo orden de numeraci\'on de las aristas que en la matriz de incidencias A , es decir, dicho de otro modo, $AC^T \equiv 0 \pmod{2}$.

Se llama **grafo euleriano** a un grafo no orientado que posee la siguiente propiedad: existe un ciclo que pasa exactamente una vez por cada una de las aristas del grafo. El ciclo en la definici\'on de grafo euleriano se denomina **euleriano**.

Teorema 8. Un grafo no orientado finito G es euleriano cuando y sólo cuando es conexo y todos los vértices suyos son de grado par.

Demostración. Es evidente que nuestras condiciones son necesarias, puesto que cada vez que el ciclo euleriano pasa por algún vértice, ha de entrar en él por una arista y salir por la otra.

A la inversa, supongamos ahora que G es conexo y todos los vértices suyos son de orden par. Convengamos en considerar un punto arbitrario x del grafo G como origen de la cadena P y prolonguémosla, cuanto sea posible, todo el tiempo a través de nuevas aristas. Por cuanto en cada vértice el número de aristas es par, el proceso puede terminar sólo en x . Si P contiene no todas las aristas del grafo G , eliminemos de G una parte de P compuesta por las aristas de este ciclo.

Los grafos P y G tienen vértices de grado par; lo mismo ha de ser válido también para el grafo restante $\bar{P}$. Por cuanto el grafo G es conexo, en $\bar{P}$ debe encontrarse un vértice y que sea incidente con las aristas de $\bar{P}$. A partir de y puede construirse una cadena nueva P' que contenga sólo aristas de $\bar{P}$. De nuevo tal cadena quedará terminada sólo cuando vuelva a y . Pero en este caso de P y P' podemos hacer un ciclo nuevo

$$P_1 = P(x, y) \cup P' \cup P(y, x),$$

el cual vuelve a x y contiene más aristas que P . Si P_1 no es un ciclo euleriano, la construcción se repite. Finalizado este proceso, el ciclo euleriano será construido.

Se denomina *ciclo hamiltoniano* (contorno hamiltoniano) a un ciclo simple (contorno simple) que pasa por todos los vértices del grafo. Un grafo se llama *hamiltoniano*, si contiene un ciclo hamiltoniano.

En las aplicaciones de los grafos a los juegos los vértices corresponden a posiciones diferentes. Así pues, la existencia del ciclo hamiltoniano es equivalente a la de una sucesión cíclica de jugadas que contiene cada posición una sola vez. De ejemplo sirve el conocido problema del caballo ajedrecístico: ¿se podrá realizar una jugada con el caballo, comenzando con un escaque arbitrario en el tablero de ajedrez, en una sucesión tal que permita pasar por cada uno de los 64 escaques y regresar al escaque de partida. En la fig. 5.12 se muestra una de las soluciones posibles.

56	41	58	35	50	39	60	33
47	44	55	40	59	34	51	38
42	57	46	49	36	53	32	61
45	48	43	54	31	62	37	52
20	5	30	63	22	11	16	13
29	64	21	4	17	14	25	10
6	19	2	27	8	23	12	15
1	28	7	18	3	26	9	24

Fig. 5.12.

En la fig. 5.13 se aduce el ejemplo de un grafo euleriano que no es

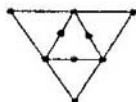


Fig. 5.13.



Fig. 5.14

hamiltoniano, y en la fig. 5.14, de un grafo hamiltoniano, mas no euleriano.

Pese a la similitud de las definiciones para los ciclos euleriano y hamiltoniano, las teorías correspondientes para estos conceptos tienen poco de común. El criterio de existencia de los ciclos eulerianos es simple (véase el teorema 8); para los ciclos hamiltonianos no se conoce una regla general. Más aún, incluso para los grafos concretos resulta, a veces, difícil decidir si es posible hallar tal ciclo. Algunas condiciones, bajo las cuales en un grafo existe el ciclo hamiltoniano, sin embargo, han sido encontradas, y se dan en ciertas obras referentes a la teoría de los grafos. Para aquellos que desean familiarizarse con estos resultados les recomendamos los libros [24], [25].

Se denomina *árbol* un grafo conexo no orientado sin ciclos. Se denomina *bosque* un grafo no orientado en el que cada componente representa un árbol.

Teorema 9. Las siguientes afirmaciones se consideran equivalentes:

- a) un grafo no orientado G es el árbol;
- b) cualesquiera dos vértices en G están unidos por una cadena única;
- c) un grafo G es conexo, mas la supresión de cualquiera de sus aristas lo hace no conexo.
- d) la adición de cualquier arista nueva al grafo G conduce a que aparezca exactamente un ciclo.

La demostración de este teorema se omite por ser trivial.

El siguiente resultado sobre los árboles pertenece a los trabajos de Cayley que investigó estos grafos en relación con las fórmulas estructurales químicas.

Teorema 10. El número de diferentes árboles que pueden ser construidos sobre n vértices dados es igual a n^{n-2} .

Demostración. Numeremos todos los vértices de un árbol T con números naturales de 1 a n . Designemos con b_1 un vértice pendiente (es decir, un vértice de grado 1) con un número mínimo, y con $l_1 = (b_1, a_1)$, la correspondiente arista pendiente. Al suprimir en T la arista l_1 y el vértice b_1 , obtendremos un nuevo árbol T_1 . Hallemos para T_1 un vértice pendiente con un número mínimo; designémoslo con b_2 , y la arista correspondiente, con $l_2 = (b_2, a_2)$. Esta reducción se repite hasta que no quede, tras la eliminación de la arista $l_{n-2} = (b_{n-2}, a_{n-2})$, la única arista $l_{n-1} = (b_{n-1}, a_{n-1})$ que une los dos vértices restantes. Entonces, la siguiente colección entre paréntesis

$$(a_1, a_2, \dots, a_{n-2})$$

llamada *código de Prufer*, se define unívocamente por el árbol T , y a dos árboles distintos T y T' les corresponden, evidentemente, diferentes códigos de Prufer.

Además, los códigos de Prufer definen los árboles T con ayuda de una construcción inversa. Si está dado el código, entonces se halla el primer vértice b_1 no contenido en él. Esto define la arista $l_1 = (b_1, a_1)$. Luego, eliminamos los vértices a_1 del código y b_1 , de la sucesión $1, 2, \dots, n$, y continuamos la construcción para los números restantes. El grafo, obtenido como resultado de esta construcción será un árbol, lo que puede ser establecido, por ejemplo, con ayuda de la inducción. Después de eliminar a_1 , el código contendrá $n - 3$ números. Si éstos corresponden al árbol T_1 , entonces el grafo T , obtenido de T_1 por adición de la arista $l_1 = (b_1, a_1)$, es también un árbol, puesto que el vértice b_1 no pertenece a T_1 . Así pues, se ha establecido que entre los códigos de Prufer y los árboles hay una correspondencia biunívoca. Mas, en el código de Prufer cada elemento puede asumir cualquiera de n valores posibles. Todos ellos corresponden a diferentes árboles. En total pueden haber n^{n-2} códigos de Prufer, de lo que precisamente se deduce nuestro teorema.

Si G es un grafo no orientado con un conjunto de vértices X , $|X| = n$, entonces, se llamará *árbol de esqueleto* (o, simplemente, esqueleto) del grafo G a todo subgrafo de esqueleto del grafo G que sea un árbol. Las aristas del grafo, pertenecientes al esqueleto, llevan el nombre de *ramas* y todas las demás aristas, se denominan *cuerdas*.

Hay situaciones cuando surge la necesidad de confeccionar una lista completa de los esqueletos del grafo G (por ejemplo, en el caso cuando hay que elegir el "mejor" árbol, mientras que el criterio que permite realizar tal elección es muy complejo, así que la resolución inmediata del problema de optimización resulta ser ejecutable). En otras situaciones, por ejemplo al hallar las funciones de transferencia de un sistema o al calcular los determinantes de ciertas matrices en la teoría macroeconómica, podemos, con ayuda de todos los esqueletos del grafo correspondiente, conseguir la simplificación de los procedimientos de cálculo.

El número de diferentes esqueletos de un grafo marcado conexo no orientado completo sobre n vértices nos da el teorema 10. Las fórmulas para el número de esqueletos en los grafos más generales pueden encontrarse en el libro [14].

He aquí uno de los resultados.

Teorema 11. Sea G un grafo no orientado sin lazos sobre n vértices, y sea A_0 su matriz de incidencias con una fila eliminada (es decir, una matriz con $n - 1$ filas independientes). Supongamos que A_0' es una matriz transpuesta respecto de la matriz A_0 . Entonces, el determinante del producto $A_0 A_0'$ es igual al número de diferentes esqueletos del grafo G , un doble (sosa) no orientado de G .

La demostración de este teorema se aduce en el libro [78]. Los algorit-

mos que generan todos los árboles de esqueleto del grafo están detalladamente analizados en [24].

Todos los árboles de esqueleto del grafo expuesto en la fig. 5.15 (hay en total 21 árboles) se aducen en la fig. 5.16. La matriz de incidencias del grafo tiene en este caso por expresión.

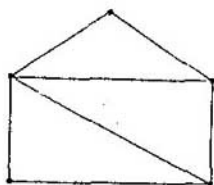


Fig. 5.15.

$$A = \begin{matrix} & \begin{matrix} a_1 & a_2 & a_3 & a_4 & a_5 & a_6 & a_7 \end{matrix} \\ \begin{matrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{matrix} & \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & -1 & -1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 & -1 & 0 & -1 \end{pmatrix} \end{matrix}.$$

Aquí consideramos que cada arista está orientada de su vértice final de índice mínimo al vértice de índice máximo.

Eliminando, por ejemplo, la fila x_2 , obtenemos la matriz A_0 . El producto de las matrices $A_0 A_0^T$ tiene en este caso la forma

$$A_0 A_0^T = \begin{pmatrix} 2 & -1 & 0 & 0 \\ -1 & 3 & -1 & 0 \\ 0 & -1 & 3 & -1 \\ 0 & 0 & -1 & 2 \end{pmatrix}.$$

El determinante del producto es igual a 21. Por consiguiente, en la fig. 5.16 se da la lista completa de los esqueletos del grafo G , expuesto en la fig. 5.15.

Muchos esqueletos obtenidos son, evidentemente, isomorfos y surge la cuestión sobre el número que hay entre ellos no isomorfos. Este es un problema más complejo, sin embargo es muy importante para muchas aplica-

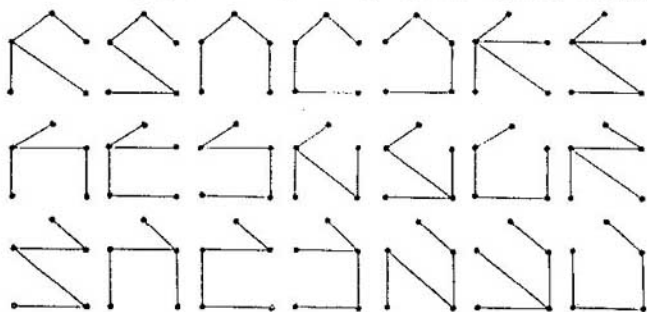


Fig. 5.16.

ciones. Es por eso que existe una literatura espaciosa también para las cuestiones análogas relacionadas con la enumeración de todos los grafos de tipos parciales. La mayoría de estas obras se apoya en la teoría de Redfield—Polya (véase § 25).

Se llama *árbol orientado* con raíz x_0 a un grafo orientado conexo, en el que del vértice x_0 a cada otro vértice conduce un camino y, además, único.

Sea G un grafo no orientado con n vértices, m aristas y p componentes conexas. El número $\varrho(G) = n - p$ se denomina *cociclomático*. El número $\nu(G) = m - \varrho(G) = m - n + p$ se denomina *ciclomático*.

En la teoría de los circuitos eléctricos los números $\varrho(G)$ y $\nu(G)$ tienen un significado físico directo. Así, por ejemplo, el número ciclomático es igual al número máximo de contornos independientes en el grafo del circuito eléctrico, es decir, al número máximo de corrientes circulares independientes que pueden fluir en el circuito. El número cociclomático es igual al número máximo de diferencias de potencial independientes entre los nudos del circuito eléctrico.

Ejercicios 9. Darse ejemplos de árboles orientados con raíces.

10. Demuéstrese que el número ciclomático de un árbol es igual a 0.

11. Demuéstrese que cada árbol es un grafo bipartido.

Sea T un esqueleto del grafo G . En el caso general, $\nu(G)$ ciclos que se obtienen por adición de cualquier cuerda de G a las ramas de T , llevan el nombre de *ciclos fundamentales*.

Cabe notar que aunque el número de ciclos fundamentales es igual a $\nu(G)$, estos propios ciclos están definidos multiformemente y dependen del esqueleto T elegido originalmente. Dicho de otro modo, el ciclo fundamental (respecto del esqueleto T) es un ciclo obtenido por adición de una cuerda al esqueleto T . Por ejemplo, para un grafo G y para su esqueleto T , representados en la fig. 5.17, los ciclos fundamentales serán:

$$\Phi_1 = \{a_1, a_{10}\}; \Phi_2 = \{a_2, a_{12}, a_{10}\}; \Phi_3 = \{a_3, a_{11}, a_{13}, a_{12}, a_{10}\};$$

$$\Phi_4 = \{a_4, a_{14}, a_{11}\}; \Phi_5 = \{a_5, a_{11}\}; \Phi_6 = \{a_6, a_{15}, a_{13}, a_{14}\};$$

$$\Phi_7 = \{a_7, a_{13}, a_{14}\}; \Phi_8 = \{a_8, a_{12}, a_{15}\}; \Phi_9 = \{a_9, a_{12}\}.$$

Si una familia de todos los ciclos fundamentales la designamos con Φ (en nuestro ejemplo $\Phi = \{\Phi_1, \dots, \Phi_9\}$), entonces cualquier otro ciclo del grafo que no pertenece a Φ puede ser expresado en forma de una combinación lineal de ciclos pertenecientes a Φ , siempre que convenimos en lo siguiente.

Supongamos que cada ciclo fundamental Φ_i ; $i = 1, 2, \dots, \nu(G)$, está representado por un vector m -dimensional en el que la j -ésima componente es igual a 1 ó a 0, lo que depende de si pertenece o no la j -ésima arista al ciclo dado. Entonces, todo ciclo C del grafo G puede representarse como suma respecto del módulo 2 de los ciclos fundamentales. Así por ejemplo, el ciclo $C_1 = \{a_3, a_{11}, a_{14}, a_6, a_8, a_1\}$ de nuestro grafo puede ser representado en la forma

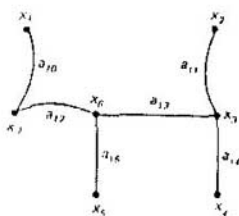
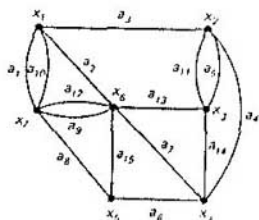


Fig. 5.17.

$$C = \Phi_1 + \Phi_3 + \Phi_6 + \Phi_8 \pmod{2} =$$

$$+ \left\{ \begin{array}{l} 10000000100000 \\ 00100000011100 \\ 000001000000111 \\ 000000010001001 \end{array} \right\} \pmod{2}.$$

$$= 101001010010010$$

Cabe notar que la inversión de la afirmación mencionada no es válida, a saber, cierta suma respecto del módulo 2 de los ciclos fundamentales no da obligatoriamente el único ciclo, pero puede representar dos y más ciclos. Por ejemplo, la suma $\Phi_2 + \Phi_3 + \Phi_6 + \Phi_7 \pmod{2}$ corresponde a dos ciclos simples $\{a_2, a_3, a_{11}, a_{13}\}$ y $\{a_6, a_7, a_{15}\}$. Así pues, para generar todos los ciclos simples del grafo G no es menester tomar todas las $2^{r(G)} - 1$ combinaciones de los ciclos fundamentales y sumarias respecto del módulo 2; algunas de estas sumas no serán, de hecho, ciclos. Más aún, si la suma dada no genera un ciclo, no se pueden rechazar otras sumas que la contienen, puesto que, sumando respecto del módulo 2 con otra suma, podemos obtener un ciclo simple.

Subrayemos también que en el grafo G podemos encontrar un conjunto de $\nu(G)$ ciclos simples independientes, que no pueden obtenerse por adición de las aristas al árbol, como lo hacíamos antes. De tal conjunto no se debe decir que es fundamental. En la fig. 5.18 se muestra el conjunto de $\nu(G) = 4$ ciclos simples independientes del grafo G , que no puede obtenerse por adición de las aristas a ningún esqueleto del grafo G y el cual, por esta razón, no es un conjunto fundamental.

Se denomina *corte* del grafo un conjunto mínimo de aristas cuya eliminación aumenta el número de componentes del grafo.

El *corte fundamental* (respecto del esqueleto T) es aquél que contiene exactamente una arista del esqueleto T . De este modo, en un grafo conexo pueden distinguirse $n - 1$ cortes fundamentales.

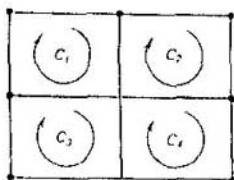


Fig. 5.18.

El teorema que sigue más abajo establece una relación entre los cortes fundamentales y los ciclos fundamentales, proporcionando el método de construcción de los cortes fundamentales.

Teorema 12. Si T es el esqueleto de un grafo no orientado G , el corte fundamental definido por la rama a_i de T está formado por a_i y aquellas cuerdas de G , las cuales, siendo adicionadas a T , proporcionan ciclos fundamentales que contienen a_i .

Demostración. Si eliminamos la arista a_i del esqueleto T , este último se descompone en dos subárboles: T_1 y T_2 . Cualquier arista, uno de cuyos vértices extremos yace en T_1 , y el otro, en T_2 , debe pertenecer al corte fundamental, puesto que la adición de cualquiera de tales aristas a las aristas de T_1 y T_2 conduce a la formación del otro esqueleto del grafo G , y, por consiguiente, cualquier conjunto privado de tales aristas no será un corte. El conjunto de estas aristas junto con la arista a_i es un corte, puesto que su eliminación divide el grafo en dos subgrafos, uno de los cuales tiene como conjunto de sus vértices T_1 , y el otro, T_2 . Esto quiere decir que este corte es fundamental. Más aún, por cuanto la arista a_i es la única por la que pasan las cadenas del esqueleto T , que tienen por origen los vértices de T_1 y termina en los vértices de T_2 las únicas aristas que cierran los ciclos fundamentales con la arista a_i incluida, serán aquellas, uno de cuyos vértices extremos yace en T_1 , y el otro en T_2 . De esta forma queda demostrado el teorema.

Sea T un esqueleto del grafo no orientado G . Se denomina *matriz de los ciclos fundamentales* Φ a una submatriz de la matriz de los ciclos C , cuyas filas corresponden a los ciclos simples, definidos por las cuerdas que están unidas con el esqueleto dado T . Dicho de otro modo, si T' es un esqueleto, entonces se llamará matriz de los ciclos fundamentales del grafo G a la matriz $\Phi = [\varphi_{ij}]$, compuesta de $\nu(G)$ filas y m columnas, en la cual

$$\varphi_{ij} = \begin{cases} 1, & \text{si la arista } a_j \text{ pertenece al ciclo fundamental } \Phi_i; \\ 0, & \text{en el caso contrario.} \end{cases}$$

Si numeramos las cuerdas del grafo G en una forma sucesiva desde 1 hasta $\nu(G)$, y las ramas, de $(\nu(G) + 1)$ a m , entonces la matriz de los ciclos fundamentales tendrá por expresión

$$\Phi = [E|\Phi_{12}],$$

donde E es una matriz unidad de orden $\nu(G)$, y Φ_{12} , una matriz de dimensión $\nu \times (m - \nu)$. Esto se explica por el hecho de que cada ciclo Φ_i contiene una y sólo una cuerda y los ciclos siempre pueden ser numerados según el número de cuerdas, a consecuencia de lo cual todas las unidades en la primera $(\nu \times \nu)$ -submatriz de la matriz Φ se hallan en la diagonal.

La *matriz de los cortes fundamentales* $K = [k_{ij}]$ se define como matriz con $n - 1$ filas y m columnas, donde

$$k_{ij} = \begin{cases} 1, & \text{si la arista } a_j \text{ pertenece al corte fundamental } k_i; \\ 0, & \text{en el caso contrario.} \end{cases}$$

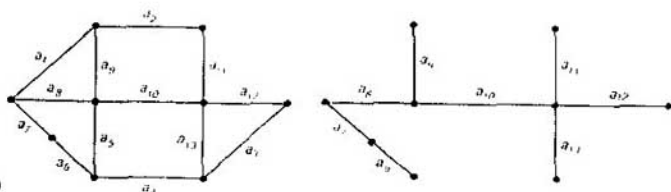


Fig. 5.19

Para la misma numeración de las aristas que en la matriz de los ciclos fundamentales Φ , la matriz K tendrá por expresión

$$K = [K_{ij}/E],$$

puesto que en este caso todo corte fundamental tiene una y sólo una rama de T .

Admitamos que el grafo G y su esqueleto T están representados en la fig. 5.19. Entonces la matriz de los ciclos fundamentales es igual a

$$\Phi = \begin{pmatrix} a_1 & a_2 & a_3 & a_4 & a_5 & a_6 & a_7 & a_8 & a_9 & a_{10} & a_{11} & a_{12} & a_{13} \\ \Phi_1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ \Phi_2 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ \Phi_3 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ \Phi_4 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ \Phi_5 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \end{pmatrix}$$

y la matriz de los cortes fundamentales K es igual a

$$K = \begin{pmatrix} a_1 & a_2 & a_3 & a_4 & a_5 & a_6 & a_7 & a_8 & a_9 & a_{10} & a_{11} & a_{12} & a_{13} \\ k_1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ k_2 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ k_3 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ k_4 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ k_5 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ k_6 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ k_7 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ k_8 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

donde los cortes k_i , correspondientes a las aristas del esqueleto, tienen los números

$$\mu(G) + i = 5 + i; \quad i = 1, 2, \dots, 8.$$

Existen varias correlaciones interesantes entre las matrices de los ciclos fundamentales y de los cortes fundamentales y las incidencias de los grafos no orientados sin lazos.

Teorema 13. La matriz de incidencias A_0 (con una fila suprimida) y la matriz transpuesta de los ciclos fundamentales Φ^T son ortogonales, es decir, $A_0 \Phi^T \equiv 0 \pmod{2}$.

Teorema 14. La matriz de los ciclos fundamentales Φ y la matriz transpuesta de los cortes fundamentales K^T son ortogonales, es decir, $\Phi \cdot K^T \equiv 0 \pmod{2}$.

Los teoremas 13 y 14 son consecuencia de dos hechos evidentes:

1) cada vértice en un ciclo es incidente al número par de aristas de este ciclo, y en el caso de un ciclo simple, a dos aristas;

2) cada corte de un ciclo, inducido por cierto corte, tiene un número par de aristas, comunes con dicho corte.

El teorema 13 se deduce del hecho 1), y el 14, del hecho 2), si recordamos que todas las operaciones se estudian con relación al módulo 2.

En lo que se refiere al teorema 14, podemos escribir

$$\Phi \cdot K^T = [L] \Phi_{12} \begin{bmatrix} K_{11} \\ F \end{bmatrix} = K_{11}^T + \Phi_{12} = 0 \pmod{2}.$$

Por eso, $K_{11}^T = -\Phi_{12} = \Phi_{12} \pmod{2}$. En otras palabras, la matriz de los cortes fundamentales puede ser obtenida en cuanto se conozca la matriz de los ciclos fundamentales, y viceversa.

Ejercicios. Cerciérese de la validez de las siguientes afirmaciones

12. Una submatriz cuadrada de dimensión $(n-1) \times (n-1)$ de la matriz de incidencias I de un grafo conexo es regular, si las columnas de la submatriz citada corresponden a las ramas de cierto árbol.

13. Supongamos que la matriz de ciclos C de un grafo conexo tiene el rango $n-m+1$. Su submatriz cuadrada B de orden $n-m+1$ es regular cuando y sólo cuando las columnas de la submatriz citada corresponden al conjunto de cuerdas respecto de cierto árbol de esqueleto del grafo dado.

14. Una submatriz regular de dimensión $(n-1) \times (n-1)$ de la matriz de incidencias I corresponde biunívocamente a los esqueletos del grafo.

15. Una matriz regular de dimensión $(n-m+1) \times (n-m+1)$ de la matriz de ciclos C corresponde biunívocamente a los complementos de los esqueletos del grafo.

16. Si las columnas de las matrices A_0 , Φ y K las escribimos valiéndonos de un mismo orden de aristas con relación a cierto árbol de esqueleto, respecto del cual están formados los ciclos y cortes fundamentales, y, a continuación, las representamos en forma de $A_0 = [I_0 | A_{12}]$, donde A_{12} es una matriz cuadrada regular de orden $n-1$, $\Phi = [L | \Phi_{12}]$ y $K = \{K_{11} | F\}$, entonces

$$K_{11} = \Phi_{12}^T = A_{12}^T A_{11}.$$

Con esto hemos de acabar con la introducción a la teoría de los grafos. Esta teoría ha acumulado un enorme material de hechos reales. A la exposición de esta teoría están dedicados toda una serie de libros. Algunos de ellos ya han sido mencionados en el texto. Sin embargo, debemos atraer la atención del lector a la obra [79] (probablemente la única de su género) sobre el origen y desarrollo de la teoría de los grafos.

CAPÍTULO 6 MÉTODOS DE RESOLUCIÓN DE LOS PROBLEMAS EXTREMALES

6.1. PLANTEAMIENTO DE LOS PROBLEMAS COMBINATORIOS EXTREMALES Y ACCESOS A SU RESOLUCIÓN

Según lo dicho en el capítulo 1, existen tres tipos de problemas combinatorios: en unos se resuelve la cuestión de existencia o no existencia de las soluciones; en otros se calcula el número de soluciones del problema; por fin, en los terceros de una colección de todas las soluciones se elige aquella que posee cierta propiedad en un grado máximo o mínimo. Los problemas de este último tipo se denominan precisamente extremales.

Demos a conocer unos cuantos ejemplos de problemas combinatorios extremales.

1. Problema de nombramientos. Se tienen n plazas vacantes y el mismo número de candidatos para ocuparlas. El nombramiento del i -ésimo candidato a la j -ésima plaza lleva consigo el pago del salario c_{ij} ; $i, j = 1, 2, \dots, n$. Para cada candidato puede ser asignada sólo una plaza, y cada plaza puede ser ocupada sólo por un candidato. En otras palabras, no se admiten el ejercer un cargo simultáneamente con otro y el fraccionamiento de los cargos. El problema consiste en llevar a cabo los nombramientos de un modo tal que la economía de los fondos de salario sea máxima, es decir, en encontrar

$$\min \sum_{i=1}^n c_{ij_i}$$

donde j_i es el cargo que obtiene el i -ésimo candidato, y $(j_1, j_2, \dots, j_n)$ es una sustitución correspondiente a los nombramientos: $i \rightarrow j_i$. Si interpretamos c_{ij} como rendimiento, el problema consistirá en la búsqueda del máximo y no del mínimo. Los problemas de nombramientos tienen muchas formas diferentes en función de las condiciones. Por ejemplo, para ejecutar n operaciones independientes están elegidos m obreros; se da una matriz $[t_{ij}]$, donde t_{ij} es el tiempo que gasta el i -ésimo obrero para realizar la j -ésima operación. La magnitud, cuyo mínimo se busca, es el tiempo de realización de todo el trabajo:

$$F = \max_i E_i; \quad E_i = \sum_{j \in R_i} t_{ij},$$

donde R_i es el conjunto de operaciones ejecutadas por el i -ésimo obrero.

En un caso particular, puede prefijarse el número de operaciones n_i para el i -ésimo obrero: $\sum_i n_i = n$. Cuando $n_i = 1$ ($i = 1, 2, \dots, n$; $n = m$), obtenemos el problema de nombramientos con el criterio de mini-máx. Por último, demos a conocer tal formulación del problema de nombramientos

cuando se toma en consideración el tiempo que gasta un obrero para pasar de una operación a la otra. Sean dadas n operaciones y m obreros. Se hacen las matrices: $T = [t_{ij}]$, donde t_{ij} es el tiempo durante el cual el i -ésimo obrero cumple la j -ésima operación, y $L = [l_{ij}]$, donde l_{ij} es el tiempo que gasta cualquier obrero para pasar de la i -ésima operación a la j -ésima. Se requiere distribuir las operaciones entre los obreros y de tal manera considerar el tiempo de paso, que el tiempo de ejecución de todo el complejo de operaciones sea mínimo.

2. Problema de un viajante de comercio. Un viajante de comercio tiene que visitar varias ciudades. Ha de elegir la ruta más corta para que empezando a viajar desde su ciudad, pueda visitar otras ciudades una sola vez y regresar. Las distancias entre las ciudades calculadas de dos en dos vienen dadas en forma de una matriz $C = [c_{ij}]$, $i, j = 1, 2, \dots, n$, donde n es el número de ciudades.

3. Problema de una mochila. Se tienen n objetos cuyos pesos son $a_1, a_2, \dots, a_n$, y el valor $s_1, s_2, \dots, s_n$, respectivamente. Se pide llenar la mochila, capaz de contener un peso no superior a R , con un surtido de objetos que posea un valor máximo.

Construyamos un vector $X = (x_1, x_2, \dots, x_n)$, donde $x_i = 1$, si el i -ésimo objeto debe colocarse en la mochila, y $x_i = 0$; en el caso contrario: $i = 1, 2, \dots, n$. El problema se formulará como problema de búsqueda de

$$\max \sum_{i=1}^n s_i x_i$$

(del valor máximo de los objetos colocados) a condición de que

$$\sum_{i=1}^n x_i a_i \leq R.$$

4. Minimización del tiempo de ejecución de un juego de operaciones. Supongamos que se pide ejecutar n operaciones. A cada operación se le hace corresponder un vértice del grafo G , donde está trazado un arco desde el vértice i hacia el j , si la operación i precede, según las condiciones tecnológicas, a la operación j . Todas las operaciones están divididas en m clases, con la particularidad de que en un momento dado puede ejecutarse sólo una operación de cada clase. Se conoce el tiempo de ejecución de las operaciones (t_j para la j -ésima operación). Se requiere determinar el orden de ejecución de todas las operaciones que asegure la realización del trabajo en el tiempo mínimo T . Por ejemplo, sea una red (esquema) de 6 operaciones: A_1, A_2, A_3, A_4 son operaciones de primera clase y B_1, B_2 , de segunda; el tiempo de ejecución de cada operación viene indicado dentro de los cuadrados en la fig. 6.1. Sean R_1 y R_2 sucesiones de ejecución de las operaciones para cada clase, respectivamente. Por ejemplo, para $R_1 = (2, 1, 3, 4)$; $R_2 = (2, 1)$ obtendremos: $T = 36$. Es evidente que el número de combinaciones admisibles depende considerablemente de la estructura de la red.

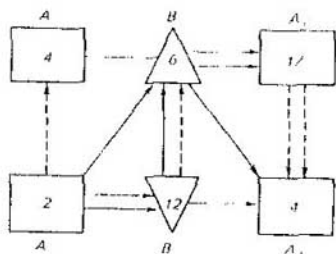


Fig. 6.1

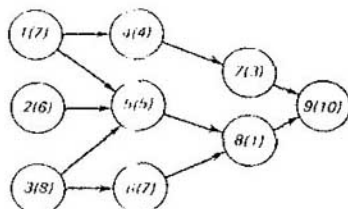


Fig. 6.2

5. Minimización del tiempo medio de mecanizado de un lote de piezas.

Se analiza un grafo orientado G con los vértices $1, 2, \dots, n$, donde el vértice i corresponde al i -ésimo tipo de las piezas a maquinar. Para cada arco (i, j) viene dada su longitud l_{ij} , la cual se interpreta como período entre el mecanizado de las piezas del i -ésimo tipo y el mecanizado de las piezas del j -ésimo tipo. La sucesión de vértices $i_1, i_2, \dots, i_k, i_{k+1}$, donde $i_{k+1} = i_1$ define el ciclo del mecanizado. Se pide minimizar el tiempo medio de este proceso

$$T_{\text{m}} = \frac{1}{k} \sum_{j=1}^k l_{i_j, i_{j+1}}$$

es decir, hallar en el grafo orientado G el contorno que tenga una longitud media mínima.

6. Minimización de la parada de una línea de montaje. Supongamos que n operaciones se efectúan en un transportador de m tramos. Para que el transportador se mueva continuamente, el tiempo de trabajo del operario no debe sobrepasar la duración del ciclo. Admitamos que en el i -ésimo tramo se realizan un conjunto de operaciones R_i . Entonces, el tiempo activo en el i -ésimo tramo es $t_i = \sum_{R_i} t_R$ y la parada (interrupción) en él es igual a $d_i = C - \sum_{R_i} t_R$, donde C es la duración total del ciclo de operaciones. La magnitud sumaria de la parada $d = \sum_{i=1}^m d_i$ debe hacerse mínima a cuenta de la distribución de las operaciones por tramos. Además, pueden plantearse las cuestiones: a) sobre el mínimo de duración del ciclo $C = \max_i T_i$ para un número dado m de tramos; b) sobre el mínimo del número m para una longitud dada del ciclo. En la fig. 6.2 vienen introducidas 9 operaciones (entre paréntesis se indica el número de unidades de tiempo que se requieren para la ejecución de la operación correspon-

diente) y 3 tramos:

$$R = (R_1, R_2, R_3) = (1, 2, 3; 4, 5, 6; 7, 8, 9),$$

tenemos

$$F_1 = 21; F_2 = 16; F_3 = 14,$$

$$C = \max_i F_i = 21; T = \sum_i t_i = 51.$$

El tiempo total de parada es

$$d = \sum_i d_i = mC - T = 3 \cdot 21 - 51 = 12.$$

7. Problema de recubrimiento. Hállese, para un grafo no orientado dado G , el recubrimiento mínimo, es decir, un surtido mínimo de aristas tal que cualquier vértice resulte ser incidente a cierta arista perteneciente al surtido mencionado. Formalicemos el planteamiento de este problema. Numeremos los vértices del grafo con los números $1, 2, \dots, m$, y las aristas, con los números $1, 2, \dots, n$. Sea $[a_{ij}]$ la matriz de incidencias del grafo G . Unamos con las aristas las variables x_j ($j = 1, 2, \dots, n$) tales que $x_j = 1$, si la arista j integra el recubrimiento; $x_j = 0$, en el caso contrario. El problema de hallar el recubrimiento mínimo es equivalente a la búsqueda del $\min \sum_{j=1}^n x_j$ con requerimientos adicionales:

$$\sum_{j=1}^n a_{ij} x_j \geq 1; i = 1, 2, \dots, m.$$

La forma general de un problema combinatorio extremal es la siguiente. Si tiene un n -conjunto de elementos, en el que se da un conjunto (finito, por regla general) de combinaciones

$$P = (\pi_1, \pi_2, \dots, \pi_r).$$

Por combinaciones $\pi_1, \pi_2, \dots, \pi_r$ pueden entenderse las permutaciones, combinaciones, diferentes sucesiones, etc. En el conjunto P se define una función F . Se necesita hallar el extremo de F (el máximo o el mínimo), o bien los elementos del conjunto P que aseguran dicho extremo.

La propia formulación de los problemas combinatorios extremales predetermina la elección de las operaciones que se aplican para su resolución. En primer lugar, hay que disponer de un conjunto, de valores de la función F y saber seleccionarlás de un modo adecuado. En segundo lugar es necesario desarrollar el método de comparar estos valores y distinguir entre ellos el valor máximo o mínimo.

La primera operación de la selección raras veces, prácticamente, resulta ser realizable, puesto que el número de toda clase de combinaciones posibles puede ser demasiado grande. Efectivamente, en el primer ejemplo el número de todas las combinaciones es igual a $n!$, en el segundo a

$(n-1)!$, en el tercer ejemplo a 2^n , en el cuarto a $(n!)^m$, en el quinto a $\sum_{k=2}^n \binom{n}{k} (k-1)$, etc.

No es más fácil la operación de comparación. En el caso general es difícil juzgar, sin realizar cálculos inmediatos, si tiene lugar $F(\pi_i) > F(\pi_j)$, o, viceversa, $F(\pi_i) \leq F(\pi_j)$. Además, la propia determinación del valor de la función $F(\pi_i)$ representa, como regla, un problema que no es nada fácil.

Las dificultades relacionadas con la elección de las variantes y la comparación de los valores son considerables. Precisamente ellas constituyen un obstáculo para el progreso de esta parte del análisis combinatorio, a pesar de su actualidad evidente. Solamente con la introducción (hace 30 años) en la práctica matemática de los ordenadores se hizo posible la resolución de toda una serie de problemas extremales.

El esquema general que caracteriza la conexión de los problemas combinatorios extremales con los métodos de programación lineal puede ser representado aproximadamente así: los elementos π_i se interpretan como puntos de un espacio euclídeo para que la función de "especial" F se haga una forma lineal. Se examina el problema de encontrar el extremo de esta función en la cápsula convexa de los puntos dados (es decir, en un poliedro convexo). En efecto, el extremo de una forma lineal en un poliedro se consigue en uno de los vértices que integran el conjunto de los elementos en consideración. Mientras tanto, el problema de hallar el extremo de una forma lineal es precisamente un problema de programación lineal. La peculiaridad de los problemas combinatorios, al realizarse tal esquema, consiste en que en el proceso de búsqueda de la solución hemos de limitarnos a los puntos cuyas coordenadas son números enteros.

Mostremos con ejemplos cómo se plantean los problemas de programación para los problemas combinatorios extremales. La solución del problema de nombramientos (ejemplo 1) representa una permutación $(p_1, p_2, \dots, p_n)$ de los números $1, 2, \dots, n$, obtenida como resultado de los nombramientos del tipo $i \rightarrow p_i$; $i = 1, 2, \dots, n$. El objetivo consiste en encontrar $\min \sum_{i=1}^n c_{ip_i}$ en el conjunto finito de las permutaciones mencionadas. Cada permutación puede interpretarse como un punto en el espacio euclídeo n^2 -dimensional; en este caso ésta puede representarse con mayor comodidad en forma de una $(n \times n)$ -matriz $X = \|x_{ij}\|$, donde $x_{ij} = 1$, si el i -ésimo candidato está designado para ocupar el j -ésimo cargo, y $x_{ij} = 0$, en el caso contrario. Sustituyamos esta condición por la otra: $x_{ij} \geq 0$, para que no se impida la aplicación de la programación. Los requerimientos referentes a la inadmisibilidad de ejercer un cargo simultáneamente con otro y al fraccionamiento de los cargos se escribirán como condiciones:

$$\sum_{j=1}^n x_{ij} = 1; \quad i = 1, 2, \dots, n;$$

$$\sum_{i=1}^n x_{ij} = 1; \quad j = 1, 2, \dots, n;$$

Los gastos sumarios, cuyo mínimo se busca, se escribirán así:

$$\sum_{i=1}^n \sum_{j=1}^n c_{ij} x_{ij}.$$

En el problema de un viajante (ejemplo 2) se busca, como sabemos,

$$\min \sum_{i=0}^n \sum_{j=0}^n c_{ij} x_{ij}$$

bajo las condiciones de que

$$\sum_{i=1}^n x_{ij} = 1; \quad j = 1, 2, \dots, n;$$

(de cada ciudad el viajante se va sólo una vez),

$$\sum_{j=1}^n x_{ij} = 1, \quad i = 1, 2, \dots, n;$$

(en cada ciudad el viajante llega una sola vez) y

$$u_i - u_j + n x_{ij} \leq n - 1; \quad i, j = 1, 2, \dots, n; \quad i \neq j.$$

La última condición está introducida con el fin de asegurar la unicidad del ciclo en el camino del viajante. Efectivamente, si existiera una solución que contenga dos o más ciclos, se encontraría un subciclo τ con k eslabones el cual no pase por el punto de partida. La sumación de las últimas condiciones para τ llevaría a una contradicción:

$$nk \leq (n - 1)k.$$

En la última condición no se han determinado u_i y u_j . Pueden hallarse para cualquier ciclo que tiene por origen el punto de partida. Si el punto i se visita en la p -ésima etapa del recorrido ($p = 1, 2, \dots, n$), entonces pongamos $u_i = p$. Para todos los i y j de aquí se deduce que $u_i - u_j \leq n - 1$. La última condición se cumple para cualesquiera $x_{ij} = 0$. En cambio, si $x_{ij} = 1$, las citadas condiciones se escriben como igualdades:

$$u_i - u_j + n x_{ij} = p - (p + 1) + n = n - 1.$$

El problema de recubrimiento (ejemplo 7) ya está formulado de tal modo que permite plantear el problema correspondiente de la programación lineal. Su planteamiento más general consiste en hallar, para un conjunto finito dado $S = \{s_1, s_2, \dots, s_n\}$ y cierta totalidad finita de sus subconjuntos S_j ($j = 1, 2, \dots, m$), un recubrimiento mínimo del conjunto S , es decir, un juego mínimo de subconjuntos S_j , en el cual todo elemento $s \in S$ pertenezca por lo menos a uno de los subconjuntos.

Para este problema general se aplica un acceso análogo: se hace una matriz de incidencias $\{a_{ij}\}$, en la cual $a_{ij} = 1$, si $S_i \in S_j$, y $a_{ij} = 0$, en el caso contrario. Hallemos x_{ij} que son iguales a uno, si S_j integra el recubrimiento, y a 0, si no lo integra. Por consiguiente, resulta aplicable al problema la programación lineal.

Para la resolución de los problemas combinatorios extremales han sido elaborados diferentes métodos de cálculo. Los más perspectivas de ellos han formado una rama aparte de la programación combinatoria (véase [80]). La idea general de dichos métodos consiste en la sustitución de la selección completa de todas las variantes por las selecciones parciales de menor volumen. Para poder realizar tal idea se hallan ciertos subconjuntos que a ciencia cierta no contienen el extremo buscado, estrechando el dominio de las variantes posibles. En esta ocasión los métodos resultan ser los más diversos y se definen por la estructura de los conjuntos finitos correspondientes. Más abajo se describen varios métodos ampliamente usados hoy día.

6.2. MÉTODO DE RAMIFICACIONES Y RESTRICCIONES¹⁾

De acuerdo con este método, el conjunto de todas las combinaciones admisibles Q se parte en subconjuntos $Q_1, Q_2, \dots, Q_{s_1}$. Cada uno de los últimos se parte, a continuación, en subconjuntos $Q_{i1}, Q_{i2}, \dots, Q_{is_2}$ ($i = 1, 2, \dots, s_1$), etc. hasta que se obtengan combinaciones separadas. Este proceso de ramificación resulta cómodo expresarlo en forma de un árbol. En cada subconjunto $Q_{i_1 i_2 \dots i_n}$ se define una función de preferencia $\varphi(Q_{i_1 i_2 \dots i_n})$. Comenzando por el vértice Q , el proceso se realiza una etapa a otra, eligiéndose cada vez un subconjunto que tenga el valor mínimo de la citada función (o el valor máximo, si lo exigen las condiciones del problema). El árbol obtenido será un árbol de soluciones.

El problema consiste en la elección acertada de la función de preferencia para que se pueda obtener una buena solución. A veces, la elección de un subconjunto se realiza al azar: en este caso la probabilidad de que se elija un conjunto dado es tanto mayor cuanto menor es el valor de su función de preferencia (funciones randomizadas de preferencia).

El problema principal en este método consiste en elegir el procedimiento de determinación de la frontera inferior (o superior). No es siempre fácil obtener el valor bastante exacto de ella, pero, conseguido este valor, el número de ramas que se examinan en el árbol de soluciones se reduce en el caso general. Esto se debe a que, si el valor de la frontera inferior para tal o cual subconjunto es superior o igual (en el caso de minimización;

¹⁾ Se conoce también como método de particiones progresivas y estimaciones, método de ramas y fronteras

inferior o igual, en el caso contrario) al valor de la función minimizada de una de las soluciones ya obtenidas, entonces, la rama correspondiente del árbol de soluciones se excluye del análisis.

Sea Q un conjunto de todas las soluciones admisibles de cierto problema: $Q = \{q_1, q_2, \dots, q_n\}$; en el conjunto de dichas soluciones está dada una función f . Se pide hallar un subconjunto $Q_n \subset Q$, en el cual la función f alcance su mínimo (o su máximo).

Supongamos que conocemos cierta frontera inferior b_0 para la función f en el conjunto Q . Admitase que por medio de un procedimiento se ha logrado partir Q en un conjunto A y un complemento de éste $\bar{A}$, y además, precisar, después, las fronteras inferiores de f en los conjuntos A y $\bar{A}$, las cuales son iguales a b_1 y b_1' , respectivamente, con la particularidad de que $b_1 \geq b_0$; $b_1' \geq b_0$. Supongamos ahora que con ayuda de otros procedimientos logramos partir Q en dos partes: B y $\bar{B}$; C y $\bar{C}$, ... Examinemos los conjuntos $A \cap B$; $A \cap \bar{B}$; $A \cap C$; $A \cap \bar{C}$ y las fronteras inferiores precisadas de f en dichos conjuntos b_2 , b_1' , b_2'' y b_1'' , respectivamente. En este caso, además,

$$\begin{aligned} b_2 &\geq b_1 \geq b_0; & b_2'' &\geq b_1' \geq b_0; \\ b_1'' &\geq b_1 \geq b_0; & b_2'' &\geq b_1' \geq b_0. \end{aligned}$$

Así pues, podemos construir un árbol orientado con la raíz Q (bisárbol) (fig. 6.3), con la particularidad de que no hay necesidad de hacerlo hasta el fin. Supongamos que hemos construido una parte del árbol partiendo el conjunto Q en ciertos subconjuntos y encontramos las fronteras inferiores para los vértices correspondientes a los subconjuntos citados. Elegimos de todos los vértices pendientes aquél que tiene una frontera mínima, y, a continuación, partiendo los subconjuntos correspondientes, obtenemos dos vértices nuevos, en los cuales precisamos las fronteras inferiores de la función f .

Enunciemos la siguiente afirmación sencilla: en cualquier etapa la unión de los subconjuntos, correspondientes a los vértices pendientes, da todo el Q . Por eso, si obtenemos, como resultado del proceso dado, un vértice compuesto por un solo conjunto $\{q_i\}$, y si la frontera inferior de la función f en el citado vértice es menor que en los restantes vértices pendientes entonces f toma en q_i su valor mínimo. El proceso se da por terminado.

Con ayuda de los razonamientos análogos se puede describir el algoritmo para hallar la solución máxima, si ésta existe.

Desgraciadamente, el método en consideración no da una respuesta constructiva al problema planteado. En cada problema concreto hay que idear algoritmos de partición del conjunto Q y los mejores procedimientos para precisar las fronteras inferiores b en los vértices pendientes.

Veamos cómo se aplica el método al problema del viajante. En el lenguaje de la teoría de los grafos en este problema se trata de la búsqueda de los contornos hamiltonianos óptimos.

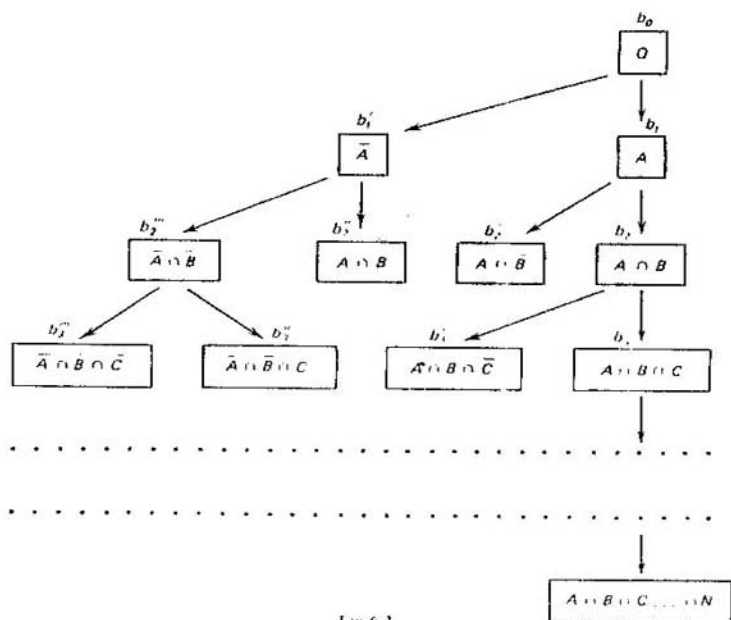


Fig. 6.3.

Sin embargo, no todos los grafos contienen el contorno de Hamilton. Por consiguiente, antes de proceder a buscar el contorno hamiltoniano óptimo, hemos de tratar, por lo menos, de establecer si tal contorno existe en el grafo dado. Por otra parte, existen grafos que disponen de una gran cantidad de contornos hamiltonianos, lo que hace imposible la selección completa incluso en los ordenadores. Por ejemplo, para un grafo completo en n vértices existen $(n - 1)!$ diferentes contornos hamiltonianos. En 1963 Little encontró un método estricto de optimización para los problemas del tipo dado con gran número de contornos hamiltonianos.

Sea dado un grafo orientado $G = (X, U)$, donde $|X| = n$, y a todo arco $(x_i, x_j) \in U$ le está asignado el valor $c_{ij} = c(x_i, x_j)$. Vamos a suponer que cualesquiera $c_{ij} \geq 0$. Si es que el arco $(x_i, x_j) \notin U$, entonces $c_{ij} = \infty$.

El algoritmo para encontrar el contorno hamiltoniano mínimo consistirá en la construcción de un bisárbol descrito en el método de ramificaciones y restricciones.

A) Q (la raíz del bisárbol) es el conjunto de todos los contornos hamiltonianos. La frontera inferior de E , es, obviamente, un cero, pero encontraremos la frontera mayor en el proceso de búsqueda de los conjuntos

A y $\bar{A}$ del bisárbol. Procedamos de la manera siguiente. En cada fila de la matriz $[c_{ij}]$ encontramos el elemento mínimo y lo restamos de todos los elementos de dicha fila. Si en la matriz obtenida hay columnas que no tienen elementos nulos, encontramos en cada una de ellas el elemento mínimo y lo restamos de todos los elementos de la columna. De resultas, tendremos una matriz nueva $[c'_{ij}]$ en la que en cada fila y en cada columna se contiene por lo menos un solo cero.

B) Sumemos todos los elementos que hemos sustraído en A). La suma de ellos b será precisamente la frontera inferior para Q . Efectivamente, tomemos un contorno arbitrario L , en el cual la suma de los valores de los arcos es igual a l . En el punto A), al transformar la matriz $[c_{ij}]$ en $[c'_{ij}]$, hemos sustraído ciertos números de todos los elementos de las filas correspondientes y de ciertas columnas. Con ello, cada vez de la magnitud l estos números se restaban también y, al fin y al cabo, la longitud del contorno L en la matriz $[c'_{ij}]$ es también no negativa. Por consiguiente, $l \geq b$.

C) Pasemos a la construcción de los vértices del bisárbol del siguiente nivel. Sea $c'_{ij} = 0$. Definamos γ_{ij} como suma del elemento mínimo de la i -ésima fila y del elemento mínimo de la j -ésima columna (excluyendo c_{ij} que es igual a cero) y hallemos

$$\theta = \gamma_{kl} - \max_{i, j \in L} \gamma_{ij}.$$

Veamos la propiedad p_{kl} : "el contorno no contiene el arco (x_i, x_j) " que se aplicará a los arcos con $c'_{ij} = 0$.

Si un contorno hamiltoniano no contiene el arco (x_i, x_j) , entonces obligatoriamente empleará ciertos dos arcos (x_i, x_r) ; $r \neq j$, y (x_r, x_j) , $r \neq i$.

D) Demostremos que $b' = b + \gamma_{kl}$ es la frontera inferior para el conjunto de contornos que satisfacen la propiedad p_{kl} . En efecto, analicemos un contorno hamiltoniano arbitrario L_1 privado del arco (x_k, x_l) . Entonces, realizada la transformación A), la longitud de este contorno en la matriz renovada será no inferior a cero. Con eso queda construido el vértice del bisárbol E_{kl} con la frontera inferior $b' = b + \theta$.

E) Construimos el vértice E_{kl} que se define por la propiedad p_{kl} : "el contorno emplea el arco (x_k, x_l) ". Suprimimos de la matriz la k -ésima fila y la l -ésima columna, sustituyéndolas por ∞ :

A) o bien el valor de c'_{ik} , si el contorno no empleaba hasta este momento los arcos que entran en el vértice x_k y salen del vértice x_l , puesto que la conexión del arco (x_l, x_k) con los arcos del contorno ya elegidos conduce a la aparición de un contorno no hamiltoniano; b) o bien el valor de c'_{lm} , si el contorno contiene, además del arco (x_k, x_l) , todos los arcos de cierto camino simple del vértice x_m al vértice x_k , puesto que la adición del arco (x_l, x_m) a los arcos del contorno ya elegidos también conduce a la aparición de un contorno no hamiltoniano; c) o bien el valor de c'_{pk} , si el contorno contiene, además del arco (x_k, x_l) , todos los arcos del camino simple que

	x_1	x_2	x_3	x_4	x_5
x_1	∞	2	3	2	7
x_2	3	∞	1	6	7
x_3	1	7	∞	3	7
x_4	2	7	3	∞	3
x_5	7	6	7	3	∞

Fig. 6.4.

llevan del vértice x_i al vértice x_n , puesto que la conexión del arco (x_p, x_k) con los arcos del contorno ya elegidos conduce a la aparición de un contorno no hamiltoniano.

F) Actuamos igual que en A) con la matriz obtenida como resultado de E).

G) Actuamos igual que en B) con la matriz obtenida como resultado de F). Al agregar la suma obtenida a la frontera para el vértice antecedente (en el primer caso, para Q), obtenemos la frontera para el vértice de E).

H) Si, como resultado de F), se obtiene una matriz de orden 1, el proceso se da por terminado. Si no, pasamos a I).

I) Entre todos los vértices pendientes del bisárbol ya construido elegimos un vértice de frontera mínima.

J) Si el vértice elegido en I) correspondía en el proceso de construcción a la propiedad p_{ij} , pasamos a C). Si no, pasamos a K).

K) Supongamos que el vértice elegido correspondía en la construcción a la propiedad p_{ij} . Sustituimos el valor en la célula (i, j) de la matriz correspondiente por ∞ . En la i -ésima fila, igual que en la j -ésima columna, encontramos un elemento mínimo y lo restamos de todos los elementos de dicha fila (columna). Pasamos luego a C).

Examinemos ahora un ejemplo numérico que muestra cómo se aplica el algoritmo de Little a la búsqueda del contorno hamiltoniano mínimo de un grafo. Sea G un grafo con 5 vértices x_1, x_2, x_3, x_4, x_5 , y a cada arco de este grafo (x_i, x_j) se le asigna un número c_{ij} ; $i, j = 1, 2, \dots, 5$. Estos valores están escritos en forma de una matriz de valores en la fig. 6.4.

Construyamos un bisárbol, descrito por el método de ramificaciones y restricciones, para este ejemplo numérico.

A) De los elementos contenidos en las filas x_1, x_2, x_3, x_4, x_5 sustraemos respectivamente sus elementos mínimos 2, 3, 1, 2, 4; además, de la columna x_5 sustraemos 1 (véase fig. 6.6 a).

B) Calculamos la suma $2 + 3 + 1 + 2 + 4 + 1 = 13$. De este modo, para el vértice Q (el conjunto de todos los contornos hamiltonianos) obtenemos una frontera: $b = 13$.

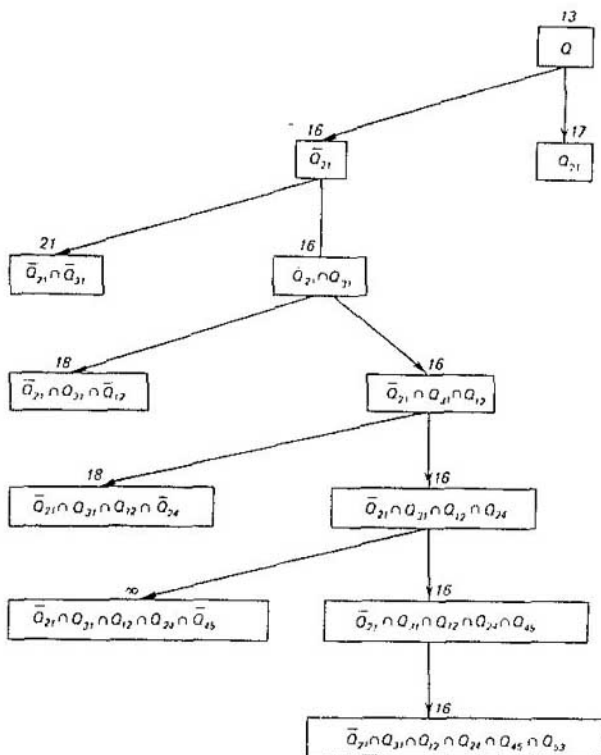


Fig. 6.5.

C) Pasamos al examen de todos los elementos nulos de la matriz:

$$\gamma(x_1, x_2) = 2; \gamma(x_1, x_4) = 2; \gamma(x_2, x_1) = 3; \gamma(x_3, x_1) = 2;$$

$$\gamma(x_4, x_3) = 0; \gamma(x_4, x_5) = 3; \gamma(x_5, x_3) = 2.$$

Por consiguiente, $\theta = \gamma(x_2, x_1) = 3$.

D) Expresamos el vértice $\bar{Q}_{21}$ del bisárbo con la frontera inferior igual a $13 + 3 = 16$.

E) Expresamos el vértice Q_{21} definido por la propiedad "el contorno contiene el arco (x_2, x_1) ". Suprimimos la segunda fila y la primera columna. En la célula (1, 2) colocamos ∞ (véase fig. 6.6 b).

F) Al principio sustraemos 2 de todos los elementos de la fila x_3 , y después, de la primera columna sustraemos 2 (véase fig. 6.6 c).

	x_1	x_2	x_3	x_4	x_5
x_1	∞	0	1	0	4
x_2	0	∞	4	3	3
x_3	0	6	∞	2	5
x_4	3	5	0	∞	0
x_5	3	2	0	2	∞

II

	x_2	x_3	x_4	x_5
x_1	∞	1	0	4
x_3	2	∞	0	3
x_4	3	0	∞	0
x_5	0	0	2	∞

a

	x_2	x_3	x_4	x_5
x_1	∞	1	0	4
x_3	6	∞	2	5
x_4	5	0	∞	0
x_5	2	0	2	∞

b

Fig. 6.6

G) La suma de los elementos sustraídos de las filas y columnas en F), es igual a 4. Por consiguiente, para el vértice $\bar{Q}_{21}$ tenemos una frontera 17 ($=13 + 4$).

H) Se ha obtenido aquí la matriz de orden 4. Pasamos a la siguiente operación I).

I) El vértice pendiente con el valor mínimo igual a 16 es $\bar{Q}_{21}$.

J) El vértice $\bar{Q}_{21}$ se ha obtenido con ayuda de la propiedad p_{21} . Pasamos a K).

K) En la célula (x_2, x_1) de la matriz expuesta en la fig. 6.6 a disponemos ∞ . Obtenemos una matriz que se muestra en la fig. 6.7 a. A continuación sustraemos 3 de todos los elementos de la fila x_2 . Véanse en la fig. 6.7 b los resultados.

C) Para la matriz expuesta en la fig. 6.7 b calculamos:

$$\gamma(x_1, x_2) = 2, \gamma(x_1, x_4) = 0, \gamma(x_2, x_4) = 0, \gamma(x_2, x_5) = 0,$$

$$\gamma(x_3, x_1) = 5, \gamma(x_4, x_3) = 0, \gamma(x_4, x_5) = 0, \gamma(x_5, x_3) = 2.$$

$$\text{Obtenemos } \Theta = \gamma(x_3, x_1) = 5.$$

D) Construimos el vértice $\bar{Q}_{21} \cap \bar{Q}_{31}$, al que corresponde la frontera $16 + 5 = 21$;

E) Construimos el vértice $\bar{Q}_{21} \cap \bar{Q}_{31}$. Suprimimos la fila x_3 y la columna x_1 . En la célula (x_1, x_2) de la matriz expuesta en la fig. 6.7 b colocamos ∞ . Véase el resultado en la fig. 6.7 c.

	x_1	x_2	x_3	x_4	x_5
x_1	∞	0	1	0	4
x_2	∞	∞	4	3	3
x_3	0	6	∞	2	5
x_4	3	5	0	∞	0
x_5	3	4	0	2	∞

a

	x_1	x_2	x_3	x_4	x_5
x_1	∞	0	1	0	4
x_2	∞	∞	1	0	0
x_3	0	6	∞	2	5
x_4	3	5	0	∞	0
x_5	3	2	0	2	∞

b

	x_2	x_3	x_4	x_5
x_1	0	∞	0	4
x_2	∞	1	0	0
x_3	5	0	∞	0
x_4	2	0	2	∞

c

Fig. 6.7

F) Todas las filas y todas las columnas de la matriz en la fig. 6.7 c contienen ceros.

G) Por consiguiente, para el vértice $Q_{21} \cap Q_{31}$ la frontera es igual a 16.

H) Se ha obtenido la matriz de orden 4. Pasamos a D).

I) El vértice pendiente de valor mínimo igual a 16 es $Q_{21} \cap Q_{31}$.

J) El vértice $Q_{21} \cap Q_{31}$ se ha obtenido con ayuda de la propiedad p_{31} . Pasamos a C).

C) Calculamos γ para todos los elementos nulos de la matriz expuesta en la fig. 6.7 c:

$$\gamma(x_1, x_2) = 2; \gamma(x_1, x_4) = 0; \gamma(x_2, x_4) = 0;$$

$$\gamma(x_2, x_5) = 0; \gamma(x_4, x_4) = 0; \gamma(x_4, x_5) = 0; \gamma(x_5, x_3) = 2;$$

$$\Theta = \gamma(x_1, x_2) = 2.$$

D) Construimos el vértice $Q_{21} \cap Q_{11} \cap \bar{Q}_{12}$ cuya frontera es 18 ($=16 + 2$).

F) Construimos el vértice $Q_{21} \cap Q_{11} \cap Q_{12}$. Suprimimos la fila x_1 y la columna x_2 de la matriz expuesta en la fig. 6.7 c. Obtenemos una matriz que se muestra en la fig. 6.8 a. En la célula (2, 3) colocamos el signo ∞ .

F) En cada fila y en cada columna de la matriz expuesta en la fig. 6.8 a se contienen ceros.

G) La frontera inferior para el vértice $\bar{Q}_{21} \cap Q_{31} \cap Q_{12}$ es igual a 26.

H) Se ha obtenido una matriz de orden 3. Pasamos a I).

	x_1	x_2	x_3
x_1	∞	0	0
x_2	0	∞	0
x_3	0	2	∞

	x_1	x_3
x_1	∞	0
x_3	0	∞

	x_1
x_1	0

Fig. 6.8.

I) El vértice pendiente de valor mínimo 16 es $Q_{21} \cap Q_{31} \cap Q_{12}$.

J) El vértice $Q_{21} \cap Q_{11} \cap Q_{13}$ se ha obtenido con ayuda de la propiedad p_{12} . Por eso, pasamos a C).

C) Calculamos γ para todos los elementos nulos de la matriz en la fig. 6.8 a:

$$\gamma(x_2, x_1) = 2; \gamma(x_2, x_3) = 0; \gamma(x_4, x_1) = 0;$$

$$\gamma(x_2, x_5) = 0; \gamma(x_5, x_1) = 2;$$

$$\Theta = \gamma(x_2, x_1) = 2.$$

D) Construimos el vértice $Q_{21} \cap Q_{31} \cap Q_{12} \cap Q_{24}$, cuya frontera será igual a 18 ($= 16 + 2$).

E) Construimos el vértice $Q_{21} \cap Q_{31} \cap Q_{12} \cap Q_{24}$. Suprimimos la fila x_2 y la columna x_3 de la matriz en la fig. 6.8 a; en la célula (4, 3) colocamos ∞ . Obtenemos una matriz que se muestra en la fig. 6.8 b.

F) En la matriz de la fig. 6.8 cada fila y cada columna contiene ceros.

G) La frontera para el vértice $Q_{21} \cap Q_{11} \cap Q_{12} \cap Q_{21}$ será igual a 16.

H) Tenemos una matriz de orden 2, por eso pasamos a I).

I) El vértice pendiente de valor mínimo 16 es $Q_{21} \cap Q_{31} \cap Q_{12} \cap Q_{24}$.

J) Este vértice se ha obtenido con ayuda de la propiedad p_{24} . Pasamos a C).

C) Calculamos γ para los elementos nulos de la matriz en la fig. 6.8 b:

$$\gamma(x_4, x_1) = \infty; \gamma(x_5, x_1) = \infty;$$

$$\Theta = \gamma(x_4, x_5) = \infty.$$

D) Construimos el vértice $\bar{Q}_{21} \cap Q_{31} \cap Q_{12} \cap Q_{45}$. La frontera para el es igual a ∞ .

E) Construimos el vértice $\bar{Q}_{21} \cap Q_{31} \cap Q_{12} \cap Q_{24} \cap Q_{45}$. Suprimimos la fila x_4 y la columna x_5 . Obtenemos una matriz que se muestra en la fig. 6.8 c.

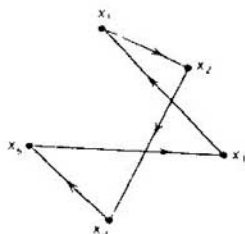
F) Es obvio que en la matriz obtenida no se requiere ninguna sustracción.

G) La frontera inferior para el vértice en E) es igual a $16 + 0 = 16$.

H) Tenemos la matriz de orden 1 con frontera mínima entre todos los vértices pendientes. Al agregar el arco (x_5, x_1) , obtenemos el contorno hamiltoniano buscado de valor 16: $(x_3, x_1, x_2, x_4, x_5)$. La solución hallada se representa en las figs. 6.9 a y 6.9 b.

	x_1	x_2	x_3	x_4	x_5
x_1	∞	2	3	2	7
x_2	3	∞	7	6	1
x_3	1	7	8	3	7
x_4	5	7	2	∞	3
x_5	7	6	4	6	∞

a



b

Fig. 6.9

El algoritmo obtenido permitió hallar la solución del problema planteado. Sin embargo, pueden existir otros contornos de Hamilton con el mismo valor. Esto resulta de que la elección de los vértices en algunos casos no era unívoca.

He aquí un ejemplo más que ilustra la aplicación del algoritmo de Little para la búsqueda en un grafo de 5 vértices del contorno hamiltoniano mínimo, si la matriz de valores está representada en la fig. 6.10 a. Continuemos el bisárbol.

A) De los elementos contenidos en las filas x_1, x_2, x_3, x_4, x_5 sustraemos respectivamente sus elementos mínimos: 3, 0, 5, 3, 16, y, a continuación, de la matriz obtenida (véase fig. 6.10 b) restamos de las columnas x_1, x_2, x_3 1, 2, 1, respectivamente (véase fig. 6.10 c).

B) Calculamos la suma $3 + 0 + 5 + 3 + 16 + 1 + 2 + 1 = 31$. Por consiguiente, para el vértice Q obtenemos la frontera inferior: $b = 31$.

C) Pasamos al examen de todos los elementos nulos de la matriz:

$$\gamma(x_1, x_3) = 17; \gamma(x_2, x_5) = 1; \gamma(x_3, x_4) = 13; \gamma(x_4, x_5) = 16;$$

$$\gamma(x_5, x_1) = 7; \gamma(x_5, x_2) = 13; \gamma(x_5, x_3) = 16; \gamma(x_5, x_4) = 1.$$

Así pues, $\Theta = \gamma(x_1, x_3) = 17$.

	x_1	x_2	x_3	x_4	x_5
x_1	∞	31	47	20	3
x_2	8	∞	17	1	0
x_3	26	20	∞	40	5
x_4	25	21	29	∞	3
x_5	17	18	17	16	∞

a

	x_1	x_2	x_3	x_4	x_5
x_1	∞	28	44	17	0
x_2	8	∞	17	1	0
x_3	21	15	∞	35	0
x_4	22	18	26	∞	0
x_5	1	2	1	0	∞

b

	x_1	x_2	x_3	x_4	x_5
x_1	∞	26	43	17	0
x_2	7	∞	16	1	0
x_3	20	13	∞	35	0
x_4	21	16	25	∞	0
x_5	0	0	0	0	∞

c

Fig. 6.10.

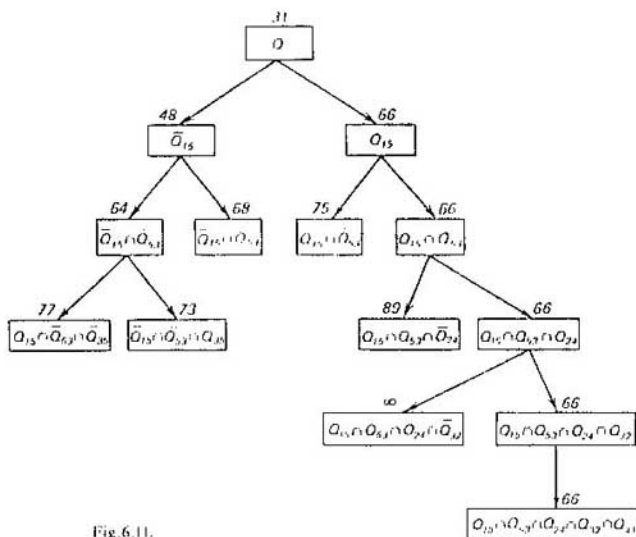


Fig. 6.11.

D) Expresamos el vértice Q_{15} del hasarbol cuya frontera inferior es igual a $31 + 17 = 48$.

E) Expresamos el vértice Q_{15} que se define por la propiedad p_{15} "el contorno hamiltoniano contiene el arco (x_1, x_5) ". Suprimimos en la matriz expuesta en la fig. 6.10 c la primera fila y la quinta columna, y en la célula $(5, 1)$ ponemos ∞ (véase fig. 6.12 a).

F) Sustraigamos de todos los elementos de las filas x_2, x_3, x_4 de la matriz obtenida, 1, 3, 16, respectivamente, y luego, de la primera columna sustraigamos además 5 (véase fig. 6.12 b).

G) La suma de los elementos sustraídos de las filas y columnas de la matriz en el punto antecedente es igual a 35. Por consiguiente, para el vértice Q_{15} la frontera inferior es igual a $31 + 35 = 66$.

H) En la última etapa se ha obtenido una matriz de cuarto orden. Por eso pasamos a la operación siguiente.

	x_1	x_2	x_3	x_4
x_1	7	∞	16	1
x_2	20	13	∞	35
x_3	21	16	25	∞
x_4	∞	0	0	0

a

	x_1	x_2	x_3	x_4
x_1	1	∞	15	0
x_2	2	0	∞	22
x_3	0	0	9	∞
x_4	∞	0	0	0

b

Fig. 6.12.

	x_1	x_2	x_3	x_4	x_5
x_1	∞	20	43	17	∞
x_2	7	∞	16	1	0
x_3	20	13	∞	35	0
x_4	21	16	25	∞	0
x_5	0	0	0	0	∞

a

	x_1	x_2	x_3	x_4	x_5
x_1	∞	9	26	0	∞
x_2	7	∞	16	1	0
x_3	20	13	∞	35	0
x_4	21	16	25	∞	0
x_5	0	0	0	0	∞

b

	x_1	x_2	x_3	x_4	x_5
x_1	∞	9	0	∞	
x_2	7	∞	1	0	
x_3	20	13	35	∞	
x_4	21	16	∞	0	

c

	x_1	x_2	x_3	x_4	x_5
x_1	∞	9	0	∞	
x_2	0	∞	1	0	
x_3	0	0	22	∞	
x_4	14	16	∞	0	

d

Fig. 6.13

I) El vértice pendiente de frontera mínima 48 es Q_{15} .

J) El vértice se ha obtenido con ayuda de la propiedad p_{15} . Pasamos a K).

K) En la célula (1, 5) de la matriz expuesta en la fig. 6.10 c colocamos ∞ . Obtenemos una matriz que se expone en la fig. 6.13 a. A continuación sustraemos 17 de todos los elementos de la fila x_1 . El resultado está representado en la fig. 6.13 b.

C) Para la matriz expuesta en la fig. 6.13 b calculamos:

$$\gamma(x_1, x_4) = 9; \gamma(x_2, x_3) = 1; \gamma(x_3, x_5) = 13; \gamma(x_4, x_5) = 16;$$

$$\gamma(x_5, x_1) = 7; \gamma(x_5, x_2) = 9; \gamma(x_5, x_3) = 16; \gamma(x_5, x_4) = 0,$$

por lo cual tenemos, por ejemplo, $\theta = \gamma(x_5, x_1) = 16$.

	x_1	x_2	x_3	x_4	x_5
x_1	∞	9	26	0	∞
x_2	7	∞	16	1	0
x_3	20	13	∞	35	0
x_4	21	16	25	∞	0
x_5	0	0	∞	0	∞

a

	x_1	x_2	x_3	x_4	x_5
x_1	∞	9	10	0	∞
x_2	7	∞	0	1	0
x_3	20	13	∞	35	0
x_4	21	16	9	∞	0
x_5	0	0	∞	0	∞

b

	x_1	x_2	x_3	x_4	x_5
x_1	∞	9	10	0	
x_2	7	∞	0	1	
x_3	12	7	0	∞	
x_4	0	0	∞	0	

c

	x_1	x_2	x_3	x_4	x_5
x_1	∞	9	10	0	
x_2	7	∞	0	1	
x_3	12	7	0	∞	
x_4	0	0	∞	0	

d

Fig. 6.14

D) Construimos el vértice $\overline{Q}_{15} \cap \overline{Q}_{53}$, al cual corresponde la frontera inferior $48 + 16 = 64$.

E) Construimos el vértice $\overline{Q}_{15} \cap \overline{Q}_{53}$. Suprimimos la quinta fila y la tercera columna en la matriz expuesta en la fig. 6.13 b, y en la célula (3, 5) colocamos ∞ . El resultado se muestra en la fig. 6.13 c).

F) De todos los elementos de la fila x_3 sustraemos primero 13, y luego de la primera columna sustraemos 7 (véase fig. 6.13 d)

G) La suma de los elementos sustraídos de las filas y columnas en F) es igual a $13 + 7 = 20$. Por consiguiente, para el vértice $\overline{Q}_{15} \cap \overline{Q}_{53}$ tenemos la frontera igual a $48 + 20 = 68$.

H) Pasamos a I).

I) El vértice pendiente $\overline{Q}_{15} \cap \overline{Q}_{53}$ tiene el valor mínimo de la frontera inferior.

J) El vértice elegido correspondía, al construir el bisárbol, a la propiedad $\overline{p}_{53}$. El valor dentro de la célula (5, 3) en la matriz expuesta en la fig. 6.13 b se sustituye por ∞ , y de todos los elementos de la columna x_3 se resta el número 16. El resultado se muestra en la fig. 6.14 b). Pasamos a C).

C) Para todos los elementos nulos de la matriz en la fig. 6.14 b calculamos:

$$\gamma(x_1, x_4) = 9; \gamma(x_2, x_3) = 9; \gamma(x_2, x_5) = 0; \gamma(x_3, x_5) = 13;$$

$$\gamma(x_4, x_5) = 9; \gamma(x_5, x_1) = 7; \gamma(x_5, x_2) = 9; \gamma(x_5, x_1) = 0.$$

De aquí, $\Theta = \gamma(x_3, x_5) = 13$.

D) Construimos el vértice $\overline{Q}_{15} \cap \overline{Q}_{51} \cap \overline{Q}_{35}$, al cual corresponde la frontera inferior $64 + 13 = 77$.

E) Para construir el vértice $\overline{Q}_{15} \cap \overline{Q}_{53} \cap \overline{Q}_{35}$ suprimimos la fila x_3 y la columna x_5 . Observemos que aquí no hay necesidad de colocar ∞ en la célula (5, 3), puesto que ella ya se encuentra allí. El resultado se muestra en la fig. 6.14 c).

F) De todos los elementos de la fila x_4 sustraemos el número 9 y obtenemos una matriz expuesta en la fig. 6.14 d).

G) Para el vértice $\overline{Q}_{15} \cap \overline{Q}_{51} \cap \overline{Q}_{35}$ tenemos una frontera inferior igual a $64 + 9 = 73$.

H) Se ha obtenido la matriz de orden 4. Por eso pasamos a I).

I) Entre los vértices pendientes del bisárbol ya construido el $\overline{Q}_{15}$ tiene frontera inferior mínima.

J) Por cuanto el vértice elegido $\overline{Q}_{15}$ correspondía durante la construcción a la propiedad p_{15} , pasamos a C).

C) Para la matriz expuesta en la fig. 6.12 b calculamos:

$$\gamma(x_5, x_4) = 0; \gamma(x_2, x_4) = 1; \gamma(x_3, x_2) = 2; \gamma(x_4, x_1) = 1;$$

$$\gamma(x_4, x_2) = 0; \gamma(x_5, x_2) = 0; \gamma(x_5, x_1) = 9. \text{ De aquí, } \Theta = \gamma(x_5, x_3) = 9.$$

D) Construimos el vértice $\overline{Q}_{15} \cap \overline{Q}_{53}$, al cual corresponde la frontera inferior $66 + 9 = 75$.

	x_1	x_2	x_3
x_1	1	∞	0
x_2	∞	0	22
x_3	0	0	∞

a

	x_1	x_2
x_1	∞	0
x_2	0	∞

b

	x_1
x_1	0

c

Fig. 6.15

E) Construimos el vértice $Q_{15} \cap Q_{53}$. Con este fin suprimimos la fila x_5 y la columna x_3 en la matriz expuesta en la fig. 6.12 b, y colocamos ∞ en la célula (3, 1), puesto que la inclusión del arco (x_3, x_1) conduce a la aparición del contorno x_1, x_5, x_3, x_1 , que no es hamiltoniano. El resultado se presenta en la fig. 6.15 a.

F) Todas las filas y columnas de la matriz expuesta en la fig. 6.15 a contienen ceros.

G) Por consiguiente, para el vértice $Q_{15} \cap Q_{53}$ tenemos una frontera inferior igual a 66.

H) Se ha obtenido la matriz de orden 3. Pasamos a I).

I) El vértice pendiente de frontera inferior mínima es $Q_{15} \cap Q_{53}$.

J) El vértice $Q_{15} \cap Q_{53}$ se ha obtenido con ayuda de la propiedad p_{53} . Pasamos a C).

C) Calculamos γ para todos los elementos nulos de la matriz expuesta en la fig. 6.15 a: $\gamma(x_2, x_4) = 23$; $\gamma(x_1, x_2) = 22$; $\gamma(x_3, x_1) = 1$; $\gamma(x_4, x_2) = 0$. De aquí $\Theta = \gamma(x_2, x_4) = 23$.

D) Construimos el vértice $Q_{15} \cap Q_{53} \cap \bar{Q}_{24}$, cuya frontera inferior es igual a $66 + 23 = 89$.

E) Construimos el vértice $Q_{15} \cap Q_{53} \cap Q_{24}$. Suprimimos en la matriz expuesta en la fig. 6.15 a la fila x_2 y la columna x_4 , sustituyendo el valor en la célula (4, 2) por ∞ . Obtenemos una matriz que se expone en la fig. 6.15 b.

F) Cada fila y cada columna de la matriz obtenida contiene ceros.

G) La frontera inferior para el vértice $Q_{15} \cap Q_{53} \cap Q_{24}$ es igual a 66.

H) Se ha obtenido la matriz de orden 2. Pasamos a I).

I) La frontera inferior mínima la tiene el vértice pendiente $Q_{15} \cap Q_{53} \cap Q_{24}$.

J) El vértice $Q_{15} \cap Q_{53} \cap Q_{24}$ se ha obtenido con ayuda de la propiedad p_{24} . Por eso, pasamos a C).

C) Calculamos: $\gamma(x_3, x_2) = \infty$; $\gamma(x_4, k_1) = \infty$. Quiere decir, $\Theta = \gamma(x_3, x_2) = \infty$.

D) Construimos el vértice $Q_{15} \cap Q_{53} \cap Q_{24} \cap \bar{Q}_{32}$. La frontera para dicho vértice es igual al infinito (∞).

E) Construimos el vértice $Q_{15} \cap Q_{53} \cap Q_{24} \cap Q_{32}$. Suprimimos la fila x_3 y la columna x_2 . Obtenemos la matriz expuesta en la fig. 6.15 c.

	x_1	x_2	x_3	x_4	x_5
x_1	∞	31	47	20	3
x_2	8	∞	17	1	0
x_3	26	20	∞	40	5
x_4	25	21	29	∞	3
x_5	17	18	17	16	∞

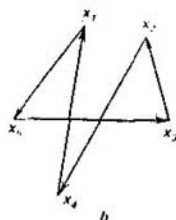


Fig. 6.16.

F) Es obvio que en la matriz obtenida no hay necesidad de sustraer nada.

G) La frontera inferior para el vértice en E) es igual a $66 + 0 = 66$.

H) La matriz obtenida es de orden 1, y el vértice $Q_{15} \cap Q_{53} \cap Q_{24} \cap Q_{32}$ tiene una frontera inferior mínima entre todos los vértices pendientes del bisárbol. Agregando el arco (x_4, x_1) , obtenemos el contorno hamiltoniano buscado de longitud 66: $(x_1, x_5, x_3, x_2, x_4, x_1)$. La solución hallada está representada en la fig. 6.16 a y b.

No es difícil modificar el algoritmo de Little para resolver los problemas referentes a la búsqueda del contorno hamiltoniano de valor máximo. No obstante, al cambiar un tanto las condiciones, reduzcamos el problema al tipo ya analizado.

Agreguemos el símbolo $(-\infty)$ a todos los pares de vértices (x_i, x_j) , para los cuales en el grafo $G = (X, U)$ no hay arco (x_i, x_j) . Sea $[c_{ij}]$ una matriz de valores del grafo $G = (X, U)$, $|X| = n$. Construyamos una matriz nueva cuyos elementos los hallamos por la fórmula

$$c'_{ij} = C - c_{ij} \quad (i, j = 1, 2, \dots, n), \text{ donde } C = \max_{i,j} c_{ij}.$$

Apliquemos ahora a la matriz $[c'_{ij}]$ el algoritmo de Little para determinar los contornos hamiltonianos mínimos. La solución mínima obtenida nos ofrece la solución máxima para la matriz $[c_{ij}]$.

6.3. MÉTODOS HEURÍSTICOS

Los métodos exactos de optimización mencionados más arriba requieren, por regla general, unos cálculos muy voluminosos. Por eso, tiene gran importancia práctica la creación de métodos sencillos que aseguran la obtención de soluciones, que son bastante próximas a las óptimas. Los métodos que no prevén la estimación de la proximidad de las soluciones obtenidas a las óptimas suelen llamarse *heurísticos*. Un método se llama *aproximado*, si admite la posibilidad de estimar la desviación de una solución de la óptima. Naturalmente, el método heurístico puede pasar al orden de aproximado tras un exitoso análisis teórico.

En los métodos heurísticos con mayor frecuencia se emplea la llamada optimización local.

Sea un conjunto de combinaciones $P = \{\pi_1, \pi_2, \dots, \pi_N\}$; se plantea el problema de calcular el mínimo de la función F , definida en este conjunto, y buscar las combinaciones sobre las cuales dicho mínimo se alcanza. El método de optimización local consiste en lo siguiente.

Para cada combinación $\pi_i \in P$ definamos un conjunto Q_i de combinaciones que se llamarán vecinas de π_i . En el lenguaje de la teoría de los grafos esto significará lo siguiente: a cada combinación π_i se le hace corresponder cierto vértice del grafo G . Los vértices vecinos π_i y π_j del grafo G se unen con arco (i, j) ($\pi_j \in Q_i$). El grafo obtenido se denominará grafo de vecindad de las combinaciones. En el grafo G puede haber una totalidad Z de sus vértices, en la cual

$$F(\pi_i) = F(\pi_j), \text{ siempre que } \pi_i, \pi_j \in Z;$$

$$F(\pi_i) < F(\pi_j), \text{ siempre que } \pi_i \in Z; \pi_j \notin Z; (i, j) \in U.$$

donde U es el conjunto de arcos del grafo de vecindad. Tal conjunto Z se llama aislado. Este conjunto puede estar compuesto, por supuesto, de un solo vértice.

El primer paso del acceso local, del cual se trata aquí, consiste en que, al elegir una combinación arbitraria π_i , hallar para ella el grafo de vecindad, después de lo cual determinar en este grafo los valores de $F(\pi_j)$ para cualquier $\pi_j \in Q_i$.

La segunda etapa consiste en una operación a la que se acostumbra llamar descenso. Encontramos $\pi_{j_0} \in Q_i$ de tal manera que

$$F(\pi_{j_0}) = \min_{\pi_j \in Q_i} F(\pi_j),$$

y si $F(\pi_{j_0}) < F(\pi_i)$, pasamos a la combinación π_{j_0} .

Procediendo de esta manera, llegaremos a un conjunto aislado tras un número finito de pasos. Sin embargo, ya encontrándonos en un conjunto aislado, no podemos estar seguros de que se ha alcanzado precisamente un extremo local (mínimo, en el caso dado), y no el general. Por eso tenemos que continuar las "pruebas" locales. Uno de los métodos de continuación de la búsqueda puede ser la elección de una nueva combinación admisible y la repetición de la operación de descenso en el antiguo grafo de vecindad. Son posibles también otros métodos. Definamos, por ejemplo, una sucesión de grafos de vecindad $G_1, G_2, \dots, G_s$ (grafos de 1-, 2-, ..., s -vecindad). Después de caer en G_1 sobre un conjunto aislado de vértices Z , pasamos al grafo G_2 partiendo de un vértice arbitrario $\pi_i \in Z$. Si Z queda aislado también en G_2 , pasamos a G_3 , etc. Si resulta posible el descenso del vértice π_i al conjunto aislado Z_1 , entonces, al pasar a algún vértice $\pi_j \in Z_1$, volveremos al grafo G_1 y repetiremos el proceso, partiendo esta vez del vértice π_j .

En el procedimiento descrito hay muchas cosas indeterminadas. Sobre

todo no está claro cómo se deben construir los grafos de vecindad y por qué dicha construcción es mejor que la elección aleatoria de las combinaciones. La experiencia muestra, no obstante, que la obtención de las combinaciones vecinas es, a veces, más fácil que construir las sucesiones aleatorias. El análisis de un problema concreto lleva frecuentemente al método de determinación de las combinaciones vecinas que asegura el descenso eficaz hacia una solución óptima o, en todo caso, bastante buena.

Veamos algunos ejemplos de determinación de la vecindad.

1. Si las combinaciones son simplemente permutaciones, entonces las combinaciones vecinas se definen como obtenidas por conmutación de los elementos próximos. Llamemos combinaciones 1-vecinas aquellas que se obtienen por conmutación de dos elementos vecinos; combinaciones 2-vecinas, aquellas que se obtienen por conmutación de dos elementos que ocupan los lugares pares o impares, etc. Por ejemplo, una permutación $\pi_1 = (1, 2, 3, 4)$ tiene cuatro combinaciones 1-vecinas: $\pi_2 = (2, 1, 3, 4)$, $\pi_3 = (1, 3, 2, 4)$, $\pi_4 = (1, 2, 4, 3)$ y $\pi_5 = (4, 2, 3, 1)$ (los elementos 1 y 4 también se consideran vecinos) y dos combinaciones 2-vecinas: $\pi_6 = (3, 2, 1, 4)$, $\pi_7 = (1, 4, 3, 2)$. Con tal definición las combinaciones 3-vecinas y 1-vecinas coinciden.

En el caso general, las combinaciones 2-vecinas se definen del modo siguiente: π_i es una combinación 2-vecina para π_j , si existe π_k tal que π_i es una combinación 1-vecina para π_k , y π_k es una combinación 1-vecina para π_j . Análogamente se introducen las combinaciones 3-vecinas, etc.

2. En los problemas 1 y 6 del § 6.1 se ha estudiado la cuestión sobre el nombramiento óptimo y sobre la parada mínima de una línea de montaje. En dichos problemas el valor de F estaba definido como $\max F_j$ para cada tramo del transportador. Los tramos, para los cuales $F_i = \max F_j$ se llamarán críticos. En tal caso resulta natural definir las combinaciones vecinas del modo siguiente: la combinación π_i es 1-vecina para π_j si puede obtenerse de π_j por conmutación de dos operaciones, una de las cuales está designada para el tramo crítico y la otra, no.

Veamos, por ejemplo, el esquema del problema 6 (fig. 6.2) una vez más. Para la solución $R^0 = (1, 2, 3; 4, 5, 6; 7, 8, 9)$ obtendremos: $F_1 = 21$; $F_2 = 16$; $F_3 = 14$; $F = 21$. Resultó ser crítico el primer tramo. Las permutaciones son admisibles sólo entre las operaciones de los tramos primero y segundo. Tenemos:

$R^1 = (1, 2, 4; 3, 5, 6; 7, 8, 9)$; $F_1 = 17$; $F_2 = 20$; $F_3 = 14$; $F = 20$;
 $R^2 = (1, 3, 4; 2, 5, 6; 7, 8, 9)$; $F_1 = 19$; $F_2 = 18$; $F_3 = 14$; $F = 19$;
 $R^3 = (1, 3, 6; 2, 4, 5; 7, 8, 9)$; $F_1 = 22$; $F_2 = 18$; $F_3 = 14$; $F = 22$;
 $R^4 = (2, 3, 6; 1, 4, 5; 7, 8, 9)$; $F_1 = 21$; $F_2 = 16$; $F_3 = 14$; $F = 21$.

En la solución R^2 el primer tramo quedó crítico pero el análisis de las soluciones vecinas no mejora la situación, pues son admisibles sólo las permuta-

ciones de los dos primeros tramos, mientras que el segundo tramo tiene el tiempo de trabajo $F_2 = 18$, próximo al tiempo crítico $F = 19$. Para disminuir F_2 , estudiemos las permutaciones admisibles de las operaciones destinadas para los tramos segundo y tercero.

Se tienen dos permutaciones admisibles de las operaciones 6, 7 y de las 5, 7. Las combinaciones correspondientes son:

$$R^5 = (1, 3, 4; 2, 5, 7; 6, 8, 9); \quad F_1 = 19; \quad F_2 = 14; \quad F_3 = 18; \quad F = 19;$$

$$R^6 = (1, 3, 4; 2, 6, 7; 5, 8, 9); \quad F_1 = 19; \quad F_2 = 16; \quad F_3 = 16; \quad F = 19.$$

En ambas soluciones, R^5 y R^6 , el valor de F_2 disminuyó. Realicemos la siguiente etapa del descenso, al tomar, por ejemplo, la solución R^6 . Las soluciones vecinas de la misma son:

$$R^7 = (1, 2, 3; 4, 6, 7; 5, 8, 9); \quad F_1 = 21; \quad F_2 = 14; \quad F_3 = 16; \quad F = 21;$$

$$R^8 = (1, 2, 4; 3, 6, 7; 5, 8, 9); \quad F_1 = 17; \quad F_2 = 18; \quad F_3 = 16; \quad F = 18;$$

$$R^9 = (1, 3, 6; 2, 4, 7; 5, 8, 9); \quad F_1 = 22; \quad F_2 = 13; \quad F_3 = 16; \quad F = 22.$$

Demostremos que la solución R^6 es óptima. Con este fin veamos la aplicación del método de ramificaciones y restricciones al problema 6b del § 6.1. Definamos el proceso de ramificaciones del modo siguiente. Supongamos predestinadas las operaciones para k tramos, es decir, están definidos $R_1, R_2, \dots, R_k$ conjuntos. Llamemos el conjunto R_{k+1} admisible, si: 1) la designación del conjunto de operaciones R_{k+1} en el $(k+1)$ -ésimo tramo no perturba la sucesión del montaje; 2) $\sum_{i \in R_{k+1}} t_i \leq C$ (C es la duración prefijada del ciclo); 3) no existe un conjunto R' tal, que $R_{k+1} \subseteq R'$,

que satisfaga las condiciones 1 y 2. La estimación desde abajo del número de tramos para cualquier subconjunto de soluciones tiene la forma: $\varphi(R_1, \dots,$

$R_k) = k + \frac{1}{C} \sum_{i \in R_k} t_i$, donde $R(k)$ es el conjunto de operaciones por ahora

no designadas.

Tomemos en nuestro ejemplo $C = 17$. Para la designación de las operaciones en el primer tramo hay cinco posibilidades:

$$R_1^1 = (1, 4, 7); \quad \varphi(R_1^1) = 1 + \frac{37}{17} = 3 \frac{3}{17};$$

$$R_1^2 = (1, 2, 4); \quad \varphi(R_1^2) = 1 + \frac{34}{17} = 3;$$

$$R_1^3 = (1, 3); \quad \varphi(R_1^3) = 1 + \frac{36}{17} = 3 \frac{2}{17};$$

$$R_1^4 = (3, 6); \quad \varphi(R_1^4) = 1 + \frac{36}{17} = 3 \frac{2}{17};$$

$$R_1^5 = (2, 3); \quad \varphi(R_1^5) = 1 + \frac{37}{17} = 3 \frac{3}{17};$$

Elegimos $R_1^2 = (1, 2, 4)$. Admitamos las siguientes designaciones para el segundo tramo:

$$R_2^1 = (3, 5; 7); \varphi(R_1^2, R_2^1) = 2 + \frac{18}{17} = 3 \frac{1}{17};$$

$$R_2^2 = (3, 6); \varphi(R_1^2, R_2^2) = 2 + \frac{19}{17} = 3 \frac{2}{17}.$$

De aquí concluimos que para $C = 17$ no son suficientes tres tramos, lo que confirma el carácter óptimo de la solución elegida R^* del problema 6a.

Ejercicio. Muéstrese que si $C = 17$, en el problema 6b es óptima la designación de las operaciones para 4 tramos. Hállense dichas designaciones.

Veamos cómo se aplica la optimización local al problema del viajante de comercio. Tenemos un conjunto $A = \{a_1, \dots, a_n\}$ de ciudades y la matriz de valores $C = [c_{ij}]$ del traslado de una ciudad a_i a otra ciudad a_j . En cada paso del algoritmo construiremos una ruta $A_r = (a_{i_1}, \dots, a_{i_r})$ a base del r -subconjunto del conjunto A ; $r = 1, 2, \dots, n$. Además consideramos que tiene lugar el traslado de la ciudad a_{i_k} a la ciudad $a_{i_{k+1}}$, $k = 1, 2, \dots, r-1$, y de la ciudad a_{i_r} a la a_{i_1} . Tomamos para $r = 1$: $A_1 = (a_i)$ (la ruta consta de una ciudad arbitrariamente elegida a_i). Si A_r está construido, buscamos en el siguiente paso la ciudad $a_i \in (A \setminus A_r)$ tal que $q(a_i, A_r) = \min_{a_i \in (A \setminus A_r)} q(a_i, A_r)$, donde con $q(a_i, A_r)$ se denota la "distancia" desde a_i hasta A_r : $q(a_i, A_r) = \min \{c_{ij} | a_j \in A_r\}$. En este caso, si $q(a_1, a_2) = C_{i_1 i_2}$ para $a_{i_k} \in A_r$, entonces se toma $A_{r+1} = (a_{i_1}, \dots, a_{i_r}, a_i, a_{i_1}, \dots, a_{i_r})$. El algoritmo construido se denomina algoritmo de inclusión de la ciudad más próxima.

Ejercicio. Muéstrese que el algoritmo de inclusión de la ciudad más próxima puede ser realizado mediante un número de operaciones proporcional a n^2 .

Teorema 1. Supongamos que la $(n \times n)$ -matriz de valores C en el problema del viajante de comercio es simétrica y satisface la desigualdad triangular: $c_{ik} + c_{kj} \geq c_{ij}$ ($i, j, k = 1, \dots, n$). Entonces

$$\frac{|I_n|}{|O_n|} < 2,$$

donde I_n es la ruta prefijada por el algoritmo de inclusión de la ciudad más próxima; O_n , la ruta óptima; $|I_n|$ y $|O_n|$ son sus valores sumarios.

Demostración. Supongamos que las ciudades $a_1, a_2, \dots, a_n$ están enumeradas de un modo tal que en el i -ésimo paso del algoritmo se le agrega a la ruta la ciudad a_i . Demostraremos el teorema estableciendo una correspondencia biunívoca entre las ciudades $a_2, \dots, a_n$ y todas las aristas en O_n (a excepción de la arista mayor) de un modo tal que el valor de inclusión de la ciudad a_k entre a_i y a_j , es decir, $c_{ik} + c_{kj} - c_{ij}$ (para que el primer paso sea correcto, tomamos $c_{ij} = 0$), no sobrepase el valor duplicado de la arista en O_n , que corresponde a a_k . Obtendremos en realidad una estimación más fuerte que la indicada en la formulación: $|I_n| \leq 2(|O_n| - l_{\max})$, donde $l_{\max}$ es la arista de mayor longitud en O_n .

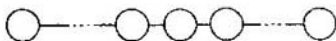


Fig. 6.17.

Estudiemos en cada paso del algoritmo no sólo la ruta corriente $a_1, \dots, a_t, 1 \leq t \leq n$, sino también algunas aristas adicionales de $O_n \setminus \{I_{\max}\}$ que unen las ciudades $a_{t+1}, a_{t+2}, \dots, a_n$ con el itinerario corriente. En el primer paso el itinerario consta de la ciudad a_1 e incluye todas las aristas en O_n , a excepción de la $I_{\max}$ (véase fig. 6.17).

Supóngase que en el $(k-1)$ -ésimo paso tenemos la ruta $a_1, \dots, a_{k-1}$ (en la fig. 6.18 se indica con una línea gruesa) con ciudades agregadas a ella $a_k, \dots, a_n$ con ayuda de las "piernas" (aristas dibujadas en línea no gruesa); una construcción de esta índole se llamará arácnica.

Admitamos que la ciudad a_k , que ha de ser incluida en la ruta en el k -ésimo paso, es la más próxima a la ciudad a_m perteneciente a la ruta (véase fig. 6.18). El valor de inclusión de a_k entre a_m y la siguiente ciudad a_t , es igual a $c_{tk} + c_{km} - c_{tm}$. Sea a_x un punto de la ruta al cual se une la pierna que incluye a_k , y sea (a_x, a_y) la primera arista de dicha pierna (puede ser que $a_y = a_k$). Por cuanto a_k es una ciudad próxima a la ruta, tenemos

$$c_{mk} \leq c_{xy}. \quad (1)$$

Según la desigualdad triangular, tenemos

$$c_{tk} \leq c_{tm} + c_{mk}. \quad (2)$$

De (1) y (2) obtenemos

$$c_{tk} \leq c_{tm} + c_{xy}. \quad (3)$$

Al adicionar (1) y (3) y al hacer uso de la simetría de la matriz C , obtenemos

$$c_{tk} + c_{km} \leq c_{tm} + 2c_{xy}.$$

lo que es equivalente a la desigualdad

$$c_{tk} + c_{km} - c_{tm} \leq 2c_{xy}.$$

es decir, la inclusión de a_k entre a_t y a_m cuesta $2c_{xy}$ como máximo. Incluida a_k , suprimimos la arista (a_x, a_y) de la configuración (la construcción sigue siendo arácnica). A la ciudad a_k se le hace corresponder la arista (a_x, a_y) . Después que el algoritmo pasa por todas las ciudades $a_2, \dots, a_n$, se establece la correspondencia requerida entre $\{a_2, \dots, a_n\}$ y $O_n \setminus \{I_{\max}\}$, lo que demuestra el teorema.

Mostremos ahora que se tiene un problema del viajante de comercio con n ciudades ($n \geq 6$), para el cual el algoritmo de inclusión de la ciudad más próxima proporciona una ruta que es dos veces mayor que la óptima. Veamos la siguiente matriz de valores C : $c_{ij} = c_{ji} = \min(j-i, n-j+i)$, es decir c_{ij} es la longitud de la ruta más corta de i a j , que va por las aristas del tipo $(k, (k+1) \pmod n)$. Para $n=8$, esta configuración viene expuesta en la fig. 6.19.

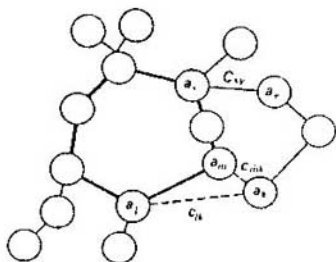


Fig. 6.18.

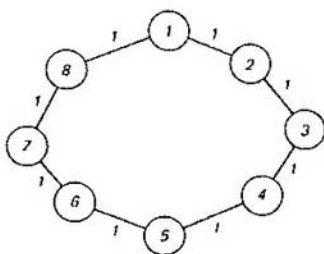


Fig. 6.19.

Demostremos que la ruta T_n compuesta por las aristas $(1, 2)$, $(n-1, n)$ y $(i, i+2)$ para $1 \leq i \leq n-2$ (en la fig. 6.20 se muestra la ruta T_8), puede obtenerse como resultado de aplicar el algoritmo de inclusión de la ciudad más próxima. Observemos que el algoritmo puede aplicarse a partir de la ciudad 1, luego se agrega a ésta la ciudad 2, la ciudad 3, etc. Sea T_k , $3 \leq k \leq n$ la ruta obtenida con ayuda de este algoritmo después de la inclusión de la ciudad k . Mostremos por inducción que T_k consta de las aristas $(1, 2)$, $(k-1, k)$ y $(i, i+2)$ para $1 \leq i \leq k-2$. Cuando $k=3$, esto es evidente. Suponiendo que esto es cierto para $3 \leq i \leq k$, observemos que T_{k+1} se obtiene de T_k por inclusión de la ciudad $k+1$ entre las ciudades $k-1$ y k .

La longitud de la ruta T_n es igual a $2n-2$, mientras que la ruta óptima O_n , compuesta por las aristas $(1, n)$ y $(i, i+1)$, $1 \leq i \leq n-1$, tiene una longitud igual a n . Por eso tenemos

$$\frac{|T_n|}{|O_n|} = 2 \left(1 - \frac{1}{n} \right).$$

Esta relación es la máxima que se admite por el teorema, puesto que $\lim_{n \rightarrow \infty} \frac{|T_n|}{|O_n|} = 1$.

La investigación teórica de la optimización local se realiza también con ayuda de los métodos teórico-probabilísticos. Con este fin en el conjunto de problemas que se resuelven por el algoritmo local dado se introduce una medida probabilística y se demuestra que con el crecimiento de la dimensión del problema, la medida de los problemas cuyas soluciones se desvían de la solución óptima en una magnitud superior a ϵ , tiende a 0, cuando la dimensión del problema crece. Tal comportamiento de las soluciones lleva el nombre de optimalidad asintótica. La optimalidad asintótica está demostrada para muchos algoritmos de solución del problema del viajante de comercio, en particular, para el más sencillo de ellos, cuando en cada paso se supone ir a la ciudad que está más próxima a aquella, en la que el viajante de comercio se encuentra en el momento dado. El lector puede encontrar una información más completa en [26].

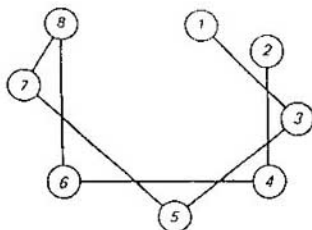


Fig.6.20.

Un tipo más de algoritmos heurísticos de optimización está relacionado con la elección aleatoria de las combinaciones conforme a cierta distribución probabilística que se elige a partir de los razonamientos de obtención más probable de las combinaciones deseables. Para datos más detallados véase [81].

"Algoritmo ávido". Veamos un conjunto finito S , a cada elemento del cual s está asignado cierto número no negativo $w(s)$, llamado peso de s . El peso del subconjunto $A \subseteq S$ se determina como suma de pesos de todos los elementos de A . Sea $\mathcal{F}$ una familia de subconjuntos del conjunto S .

Varios problemas de la optimización combinatoria se reducen al problema siguiente: hallar en la familia $\mathcal{F}$ un subconjunto de peso máximo. Por ejemplo, a este tipo de problemas se reduce el problema de hallar el esqueleto (o el bosque) de peso máximo en un grafo ponderado G , si $\mathcal{F}$ es un juego de todos los esqueletos (bosques) del grafo G .

Para resolver este problema resulta natural aplicar el siguiente algoritmo que se denomina "ávido" (greedy, en inglés). Este es como sigue:

Paso 1. Se elige un elemento s_1 tal que $\{s_1\} \in \mathcal{F}$ y $w(s_1) \geq w(s)$ para todos aquellos s que $\{s\} \in \mathcal{F}$. Si tal s_1 no existe, entonces tenemos una parada (indisponibilidad).

Paso 2. Se elige un elemento s_2 tal que $\{s_1, s_2\} \in \mathcal{F}$ y $w(s_2) \geq w(s)$ para todos aquellos $s \neq s_1$ que $\{s_1, s_2\} \in \mathcal{F}$. Si tal s_2 no existe, entonces tenemos una parada.

Paso k . Se elige tal elemento s_k , distinto de $s_1, s_2, \dots, s_{k-1}$, que $\{s_1, s_2, \dots, s_{k-1}, s_k\} \in \mathcal{F}$ y $w(s_k)$ sea máximo entre todos los s de este tipo. Si tal s_k no existe, tenemos una parada.

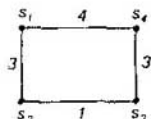
Es obvio que el algoritmo "ávido" finaliza su trabajo al obtener un subconjunto de la familia $\mathcal{F}$ que sea máximo por inclusión. Sin embargo, dicho subconjunto puede tener en $\mathcal{F}$ un peso no máximo. En efecto, si $S = \{a, b, c, d\}$, $w(a) = 4$, $w(b) = 3$, $w(c) = 2$, $w(d) = 2$, y $\mathcal{F} = \{\{a\}, \{a, c\}, \{b, c, d\}, \{b, d\}\}$, entonces, sirviéndonos del algoritmo "ávido", obtendremos el subconjunto $\{a, c\}$ cuyo peso es igual a 6, aunque el subconjunto de peso máximo es $\{b, c, d\}$ y el peso de éste es igual a 7. Al mismo tiempo, si cambiamos los pesos, poniendo $w(a) = 6$, $w(b) = 3$, $w(c) = 2$, $w(d) = 2$, el algoritmo "ávido" nos ofrece de nuevo el subconjunto $\{a, c\}$ de peso 9, que esta vez ya es un subconjunto de peso máximo.

El algoritmo "ávido" merece, obviamente, su nombre. Tiene una constante aspiración de incluir en un subconjunto de la familia $\mathcal{F}$ el elemento de peso máximo posible. No lo hace sólo en el caso en que, al agregar tal elemento, el subconjunto se hace inadmisible, es decir, no pertenecerá a $\mathcal{F}$.

Estudiamos la correlación existente entre la solución obtenida por el algoritmo "ávido" y la estructura de la familia $\mathcal{F}$.

Un par $(S, \mathcal{F})$ se llama *sistema de conjuntos*, si partiendo de que $A \in \mathcal{F}$, $A' \subseteq A$ se deduce que $A' \in \mathcal{F}$. Los elementos de la familia $\mathcal{F}$ se denominan en este caso *conjuntos independientes*.

Se puede comprobar que con ayuda del algoritmo "ávido" se puede hallar en un grafo ponderado el bosque de peso máximo. Al mismo tiempo, el algoritmo "ávido" no trabaja para el problema sobre una combinación de pares máxima para el grafo. En efecto, el algoritmo "ávido", aplicado a un grafo ponderado G :



daría la combinación de pares $\{(s_1, s_4), (s_2, s_3)\}$, mientras que la combinación de pares máxima ponderada en G es $\{(s_1, s_2), (s_3, s_4)\}$.

El análisis de los sistemas de conjuntos que son distintos por su naturaleza y origen, pero que poseen la propiedad de que el algoritmo "ávido" resuelve los correspondientes problemas combinatorios de la optimización nos conduce a la siguiente definición.

Un sistema de conjuntos $M = (S, \mathcal{F})$ se denomina *matroide*, si el algoritmo "ávido" resuelve correctamente cualquier problema combinatorio concreto referente a la optimización para el sistema M .

Teorema 2. Sea $M = (S, \mathcal{F})$ un sistema de conjuntos. Entonces son equivalentes las siguientes afirmaciones:

- 1) M es un matroide;
- 2) si $F_1, F_2 \in \mathcal{F}$ donde $|F_1| = p$, y $|F_2| = p + 1$, existe un elemento $s \in F_2 - F_1$ tal que $F_1 \cup \{s\} \in \mathcal{F}$;
- 3) si A_1 y A_2 son subconjuntos independientes, máximos por inclusión, del conjunto A , entonces $|A_1| = |A_2|$.

Demostración. (1) $\Rightarrow$ (2). Sea M un matroide, pero se supone que la afirmación 2 del teorema no se cumple, es decir, existen dos subconjuntos independientes $F_1, F_2 \in \mathcal{F}$ tales que $|F_1| = p$, $|F_2| = p + 1$, y el subconjunto $F_1 \cup \{s\} \notin \mathcal{F}$, cualquiera que sea $s \in F_2 - F_1$. Supongamos que la función ponderal sobre el conjunto S viene dada del modo siguiente:

$$w(s) = \begin{cases} p + 2, & \text{si } s \in F_1; \\ p + 1, & \text{si } s \in F_2 - F_1; \\ 0, & \text{si } s \notin F_1 \cup F_2. \end{cases}$$

Entonces, el conjunto independiente F_1 no es óptimo, es decir, no posee máximo. Efectivamente, $w(F_2) \geq (p+1)^2 > p(p+2) = w(F_1)$. El algoritmo "ávido" para este problema concreto empezará por la elección de todos los elementos del conjunto F_1 , puesto que estos elementos son de peso máximo. A continuación, el algoritmo "ávido" no podrá aumentar el peso total, ya que para todos los elementos restantes s o bien $F_1 \cup \{s\} \in \mathcal{F}$ si $s \in F_2$, o bien $w(s) = 0$. Por consiguiente, el algoritmo "ávido" da una solución no óptima de F_1 , por lo cual M no es un matroide, lo que contradice nuestra suposición. Por lo tanto, la afirmación 2 se cumple y la implicación en consideración queda demostrada.

(2) $\Rightarrow$ (3). Supongamos que la afirmación 2 del teorema se cumple y sean A_1 y A_2 dos subconjuntos independientes, máximos por inclusión, del conjunto $A \subseteq S$. Admitamos que $|A_1| < |A_2|$. Rechacemos $|A_2| < |A_1| - 1$ elementos del conjunto A_2 . Entonces, por ser el sistema $\mathcal{F}$ cerrado respecto de la inclusión, obtendremos un subconjunto independiente $A_3 \subseteq A$ tal que $|A_3| = |A_1| + 1$. En virtud de la afirmación 2 del teorema, existe un elemento $s \in A_3 - A_1$ tal que $A_1 \cup \{s\} \in \mathcal{F}$. Por consiguiente, el subconjunto A_1 no es un subconjunto independiente, máximo por inclusión, del conjunto A . La contradicción obtenida demuestra la implicación requerida.

(3) $\Rightarrow$ (1) Supongamos que tiene lugar la afirmación 3 del teorema. Mostremos que en este caso el algoritmo "ávido" resuelve el problema de la optimización combinatoria sobre el sistema de conjuntos M . Admitamos que esto no es así, es decir, para cierto conjunto de pesos $w(s)$, donde $s \in S$, el algoritmo "ávido" lleva a un conjunto independiente $F = \{s_1, s_2, \dots, s_i\}$, mientras que existe un conjunto $F' = \{s'_1, s'_2, \dots, s'_j\}$ tal que $w(F') > w(F)$. Sin restringir la generalidad de los razonamientos podemos considerar que los elementos de los conjuntos F y F' están ordenados de tal modo que $w(s_1) \geq w(s_2) \geq \dots \geq w(s_i)$ y $w(s'_1) \geq w(s'_2) \geq \dots \geq w(s'_j)$. Podemos considerar evidentemente, que F' es un subconjunto independiente, máximo por inclusión, del conjunto S . Según la construcción, el subconjunto F' es también subconjunto del conjunto S , máximo por inclusión. Por eso, en virtud de la afirmación 3 del teorema, para $A = S$ obtenemos que $i = j$. Demostremos que para todos los $m = 1, 2, \dots, i$ se verifica la desigualdad $w(s_m) \geq w(s'_m)$ lo que contradirá la suposición de que $w(F') > w(F)$, con lo que quedará por terminada la demostración de nuestra implicación.

Realicemos la demostración por inducción respecto de m . Para $m = 1$ la afirmación es evidente. Supongamos ahora que $w(s_m) < w(s'_m)$ para cierto $m > 1$, y que $w(s_p) \geq w(s'_p)$ para $p = 1, 2, \dots, m-1$. Veamos un conjunto $A = \{s \in S: w(s) \geq w(s'_m)\}$. El conjunto $\{s_1, \dots, s_{m-1}\}$ es subconjunto independiente, máximo por inclusión, del conjunto A , puesto que si $\{s_1, s_2, \dots, s_{m-1}, s\} \in \mathcal{F}$, y $w(s) \geq w(s'_m) > w(s_m)$, entonces el algoritmo "ávido" debería elegir s , en lugar de s_m , a título del siguiente elemento del conjunto F_1 . Esto contradice la afirmación 3 del teorema, por cuanto

$\{s'_1, s'_2, \dots, s'_m\}$ es otro subconjunto independiente del conjunto A de mayor potencia. Con esto termina la inducción y toda la demostración del teorema.

En el teorema 2 se ha demostrado en realidad la equivalencia existente entre el concepto de matroide (introducido aquí) como sistema de conjuntos sobre los cuales el algoritmo "ávido" da soluciones óptimas y otras definiciones de los matroides que se consideran en el capítulo 8. En el párrafo 8.4 se aducirán diferentes ejemplos de los matroides, sobre los cuales el algoritmo "ávido" es siempre eficaz.

6.4. OPTIMIZACIÓN EN LOS GRAFOS

En los párrafos antecedentes de este capítulo ya se han examinado problemas combinatorios extremales, relacionados con la optimización en los grafos (véase el problema del viajante de comercio). En el presente párrafo estudiaremos una serie más de problemas del mismo tipo.

Al principio nos dedicaremos al estudio de los grafos orientados $G = (X, \Gamma)$ que no tienen contornos. Introduzcamos para ellos la noción de función ordinal. Así se denomina una función $O(x_i)$ definida sobre un conjunto de vértices $X = \{x_i\}$, que toma valores de números enteros. En este caso, si $x_j \in \Gamma(x_i)$, entonces $O(x_j) > O(x_i)$. Dicho de otro modo, si el vértice x_j sigue tras el vértice x_i , entonces el valor de la función ordinal en el vértice x_j es superior a su valor en el vértice x_i . La magnitud $O(x_i)$ se llamará nivel del vértice x_i .

Las funciones ordinales para un grafo concreto $G = (X, \Gamma)$ pueden construirse mediante métodos diferentes. He aquí uno de éstos. Veamos los subconjuntos de los vértices $N_0, N_1, \dots, N_r$ definidos del modo siguiente:

$$\begin{aligned} N_0 &= \{x_i | x_i \in E: \Gamma^{-1}(x_i) = \emptyset\}; \\ N_1 &= \{x_i | x_i \in E \setminus N_0; \Gamma^{-1}(x_i) \in N_0\}; \\ &\dots\dots\dots \\ N_r &= \{x_i | x_i \in (E \setminus \bigcup_{k=0}^{r-1} N_k); \Gamma^{-1}(x_i) \in \bigcup_{k=0}^{r-1} N_k\}. \end{aligned}$$

Es fácil ver que los conjuntos $N_0, N_1, \dots, N_r$ forman una partición X , y la función O , que toma el valor k sobre el conjunto N_k ($k = 0, 1, \dots, r$) es una función ordinal del grafo G . En la fig. 6.21 $\Gamma^{-1}(N_r) = \emptyset$, está expuesto un grafo cuyo conjunto de vértices está partido en subconjuntos $N_0, N_1, \dots, N_r$.

A la par con la función ordinal sobre el conjunto de vértices del grafo se define también una función numérica φ . Por medio de esta función se determinan las características numéricas de los problemas, que han de ser optimizadas. Supongamos que $(x_1, x_2, \dots, x_n)$ es un camino en el grafo G , es decir, $x_1, x_2, \dots, x_n \in X; x_{i+1} \in \Gamma(x_i)$ ($i = 1, 2, \dots, n-1$). Llamemos

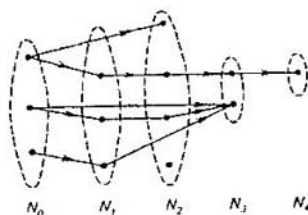


Fig. 6.21.

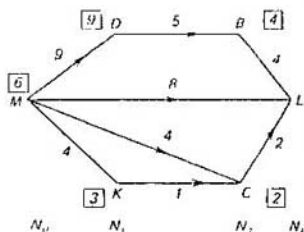


Fig. 6.22.

valor de este camino a un número $\lambda = \varphi(x_1) + \varphi(x_2) + \dots + \varphi(x_n)$.

A veces la función numérica φ se define no sobre los vértices, sino sobre los arcos. En tal caso se llama valor del camino $(x_1, x_2, \dots, x_n)$ a un número $\lambda = \varphi(x_1, x_2) + \varphi(x_2, x_3) + \dots + \varphi(x_{n-1}, x_n)$.

Muchos problemas de optimización tienen la siguiente formulación: hallar entre cierto conjunto de caminos P el mínimo (o el máximo), es decir, un camino que posea el valor mínimo (o el máximo) de λ . Como conjunto P podemos elegir, por ejemplo, el conjunto de todos los caminos hamiltonianos. En esta sección analizaremos el problema de hallar un camino mínimo y máximo entre todas los caminos que unen dos vértices fijos.

Teorema 3 (de optimalidad). Supongamos que un cierto camino que une el vértice x de nivel m y el vértice x' de nivel s , es mínimo (máximo). Entonces, su subcamino entre el vértice y de nivel k y el vértice y' de nivel p ($m \leq k < p \leq s$) es también mínimo (máximo).

Demostración. Sea el camino $(x, \dots, y, \dots, y', \dots, x')$ mínimo. Si su subcamino $(y, \dots, y')$ no es mínimo, existe, entonces, otro camino de y a y' de menor valor. Sustituyamos el subcamino anterior en el camino $(x, \dots, x')$ por uno nuevo, y obtendremos, como resultado, el nuevo camino desde x hasta x' de menor valor, lo que lleva a una contradicción.

El teorema alegado yace en la base del método para hallar los caminos mínimos en un grafo sin contornos. Este método fue propuesto por Bellman y obtuvo el nombre de método de programación dinámica.

Explicemos la esencia de este método con un ejemplo. Sea dado un grafo $G = (X, U)$ (fig. 6.22); la función numérica viene dada en los arcos. Buscaremos el camino de M a A con el valor mínimo.

La función ordinal $O(x)$ adquiere los siguientes valores:

$$O(M) = 0; O(D) = O(K) = 1; O(B) = O(C) = 2; O(A) = 3.$$

Examinamos sucesivamente, a partir de A , todos los vértices del grafo en el orden de decrecimiento de su función ordinal y asignamos a todo vértice un número igual al valor mínimo del camino que separa dicho vértice de A .

Para buscar los caminos mínimos en los grafos que tienen contornos

existen también métodos diferentes. Puede ocurrir que en tales grafos no haya caminos máximos.

Algoritmo de Ford. Supongamos que una función numérica l viene dada sobre los arcos del grafo G que no tiene contornos. Podemos considerar que l toma sólo valores positivos (de lo contrario, esto puede conseguirse con facilidad agregando un número suficientemente grande a todos los $l(x_i, x_j)$). Se requiere hallar el camino mínimo que lleva desde x_0 hasta x_n . A todo vértice x_i le asignaremos símbolos de acuerdo con el siguiente algoritmo:

1) pongamos $\lambda_i = \infty$ en cada vértice x_i , salvo $\lambda_0 = 0$;

2) busquemos un arco (x_i, x_j) tal que $\lambda_j - \lambda_i > l(x_i, x_j)$ y sustituimos λ_j por $\lambda_i + l(x_i, x_j) < \lambda_j$; procedemos de este modo hasta que se haga posible encontrar un arco que disminuya por lo menos un solo valor.

Demostremos que el número λ_n , hallado según la regla citada, será el valor del camino mínimo de x_0 a x_n (si $\lambda_n = \infty$, tal camino no existe).

Se puede indicar un vértice x_{r_1} tal que con ayuda del arco (x_{r_1}, x_n) se haya disminuido el valor de λ_n por la última vez, y

$$\lambda_n = \lambda_{r_1} + l(x_{r_1}, x_n).$$

Análogamente, existe un vértice x_{r_2} tal que

$$\lambda_{r_1} = \lambda_{r_2} + l(x_{r_2}, x_{r_1}),$$

etc. La sucesión $\lambda_n, \lambda_{r_1}, \lambda_{r_2}, \dots$ es estrictamente decreciente. Por consiguiente, para cierto k obtenemos $\lambda_{r_{k+1}} = 0$, es decir, $x_{r_{k+1}} = x_0$.

Demostremos que el itinerario $(x_0, x_{r_1}, x_{r_2}, \dots, x_n)$ es mínimo (su valor es igual a λ_n). Tomemos un camino arbitrario que lleva de x_0 a x_n : $(x_0, x_{v_1}, x_{v_2}, \dots, x_{v_m}, x_n)$. Son válidas las siguientes correlaciones

$$\lambda_{r_1} - 0 \leq l(x_0, x_{v_1});$$

$$\lambda_{r_1} - \lambda_{v_1} \leq l(x_{v_1}, x_{v_2});$$

$$\dots\dots\dots$$

$$\lambda_n - \lambda_{v_m} \leq l(x_{v_m}, x_n).$$

Al sumar término a término estas desigualdades, obtendremos

$$\lambda_n \leq l(x_0, x_{v_1}) + l(x_{v_1}, x_{v_2}) + \dots + l(x_{v_m}, x_n),$$

lo que se trataba de demostrar.

Observemos que el algoritmo de Ford puede emplearse para buscar el camino máximo en los grafos sin contornos. En este caso se debe poner $\lambda_i = 0$ para todos los vértices sin excepción alguna, y luego sustituir λ_j por $\lambda_j' = \lambda_i + l(x_i, x_j)$, si $\lambda_j' > \lambda_j$, hasta que se pueda aumentar λ_j .

Mencionemos, por fin, el algoritmo de Bellman—Kalaba. Este último emplea el llamado principio de optimalidad: "cualquier subcamino del camino mínimo es un camino mínimo entre los vértices correspondientes".

Sea un grafo orientado $G = (X, U)$; $|X| = n + 1$. Numeremos todos los vértices de 0 hasta n .

Sea l_{ij} un valor que se le asigna al arco (x_i, x_j) ($l_{ij} = \infty$, si $(x_i, x_j) \notin U$; $l_{ij} = 0$, si $i = j$). Buscaremos un camino $(x_0, x_{i_1}, \dots, x_{i_n}, x_n)$ tal que el valor del camino $l_{0,i_1} + l_{i_1,i_2} + \dots + l_{i_n,n}$ sea mínimo.

Examinemos el sistema de ecuaciones

$$v_i = \min_{j \neq i} (v_j + l_{ij}); \quad i = 0, 1, 2, \dots, n-1;$$

$$j = 0, 1, 2, \dots, n.$$

Entonces, las soluciones v_i de este sistema son valores de los caminos mínimos de x_i a x_n ; $i = 0, 1, \dots, n-1$. El hecho de que los valores de los caminos mínimos satisfacen el sistema mencionado se deduce del principio de optimalidad. La afirmación recíproca de que las soluciones del sistema son valores de los caminos óptimos se deducirá del algoritmo que se usa para hallar v_i .

Pongamos

$$v_i^{(0)} = l_{in}; \quad i = 0, 1, 2, \dots, n-1; \quad v_n^{(0)} = 0.$$

Calculamos sucesivamente:

$$v_i^{(1)} = \min_{j \neq i} (v_j^{(0)} + l_{ij}); \quad i, j = 0, 1, 2, \dots, n-1;$$

$$v_i^{(k)} = \min_{j \neq i} (v_j^{(k-1)} + l_{ij});$$

$$v_n^{(k)} = 0.$$

Los cálculos se realizan hasta que se cumpla el sistema de igualdades

$$v_i^{(k)} = v_i^{(k-1)}; \quad i = 0, 1, 2, \dots, n.$$

Del procedimiento que se usa para el cálculo se deduce que $v_i^{(m)}$ es el camino mínimo del vértice x_i al vértice x_n entre todos los caminos que pasan no más que por m arcos. Por consiguiente, para hallar la solución del sistema son suficientes $n-1$ iteraciones.

Ha de notarse que el algoritmo de Bellman—Kalaba puede ser aplicado para la búsqueda del camino máximo en un grafo sin contornos. Para ello es suficiente sólo sustituir en todas las correlaciones mín por máx y poner l_{ij} igual a $-\infty$, si el arco (x_i, x_j) no se contiene en el grafo.

6.5. FLUJOS EN LAS REDES

Al resolver los problemas extremales se emplea a menudo su interpretación en términos de los flujos en las redes. Se denomina *red* a un grafo orientado conexo $G = (X, U)$ con un conjunto de vértices X y un conjunto de arcos U , donde a todo arco $(x_i, x_j) \in U$ se le asigna una característica numérica no negativa $c(x_i, x_j)$ que se llama *capacidad de paso*. En G se distinguen dos vértices fijos: s y t ; s se denomina *entrada* (también se conoce como fuente), t se denomina *salida* (también se llama, a veces, vertedero) y los restantes vértices son *intermedios*.

Se llama *flujo estacionario* de valor v desde s a t de la red $G = (X, U) = (X, \Gamma)$ a una función numérica f , definida en el conjunto de arcos U , que satisface las siguientes ecuaciones lineales y desigualdades:

$$\sum_{y \in \Gamma(x)} f(x, y) - \sum_{y \in \Gamma^{-1}(x)} f(y, x) = \begin{cases} v, & x = s; \\ 0, & x \neq s, t; \\ -v, & x = t; \end{cases} \quad (1)$$

$$0 \leq f(x, y) \leq c(x, y) \text{ para cualesquiera } (x, y) \in U. \quad (2)$$

Introduzcamos las siguientes designaciones con el fin de simplificar las notaciones de las correlaciones de flujo.

Sean X, Y ciertos conjuntos; entonces, por (X, Y) se entenderá el conjunto de todos los pares (x, y) ; $x \in X, y \in Y$.

Pongamos para cierta función $g(x, y)$ y $h(x)$:

$$g(X, Y) = \sum_{(x, y) \in (X, Y)} g(x, y);$$

$$h(X) = \sum_{x \in X} h(x).$$

El problema del flujo máximo estacionario es el problema de maximización de la variable v que obedece a las restricciones de flujo (1) y (2).

Se denomina *corte* L en la red (X, U) , que separa los vértices s y t , a un conjunto de arcos $(Y, \bar{Y})$, donde $Y \subseteq X, \bar{Y} = X - Y, s \in Y, t \in \bar{Y}$.

El número $c(Y, \bar{Y})$ lleva el nombre de *capacidad de paso del corte* $(Y, \bar{Y})$.

Señalemos que cualquier camino de s a t contiene al menos un arco del corte $(Y, \bar{Y})$. Por eso, si de la red se excluyen los arcos de tal o cual corte, en la red nueva no quedará ningún camino de s a t , y el valor del flujo máximo por la citada red será igual a cero. Se pone claro intuitivamente que el valor v de un flujo arbitrario f no sobrepasa la capacidad de paso de cualquier corte. Demostremos que esta afirmación es cierta.

Lema 1. Sea f un flujo de s a t en la red (X, Γ) de valor v ; supongamos que $(Y, \bar{Y})$ es un corte que separa s de t . Entonces

$$v = f(Y, \bar{Y}) - f(\bar{Y}, Y) \leq c(Y, \bar{Y}).$$

Demostración. Reescribamos las relaciones (1) en una forma más compacta:

$$f(s, X) - f(X, s) = v;$$

$$f(x, X) - f(X, x) = 0; \quad x \neq s, t;$$

$$f(t, X) - f(X, t) = -v.$$

Sumemos término a término aquellas de estas relaciones que se refieren a $x \in Y$. Teniendo en cuenta que $s \in Y$, y $t \in \bar{Y}$, obtenemos

$$v = f(Y, X) - f(X, Y);$$

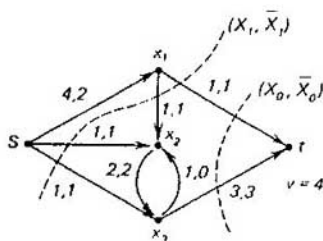


Fig. 6.23.

haciendo uso de la relación $X = Y \cup \bar{Y}$, obtenemos

$$\begin{aligned} v &= f(Y, Y \cup \bar{Y}) - f(Y \cup Y, Y) = \\ &= f(Y, Y) + f(Y, \bar{Y}) - f(Y, Y) - f(\bar{Y}, Y) = \\ &= f(Y, Y) - f(\bar{Y}, Y). \end{aligned}$$

Resta por observar que $f(\bar{Y}, Y) \geq 0$; $f(Y, \bar{Y}) \leq c(Y, \bar{Y})$, con lo que queda demostrado el lema.

El corte de capacidad de paso mínima (el cual siempre existe, puesto que el conjunto de todos los cortes es finito) se denomina mínimo. De lo demostrado más arriba se deduce que la intensidad del flujo máximo no sobrepasa la capacidad de paso del corte mínimo.

Veamos la red expuesta en la fig. 6.23, donde el primer número en el arco es igual a su capacidad de paso mientras que el segundo número, al valor de la función v . Entonces, $(X_0, \bar{X}_0)$ y $(X_1, \bar{X}_1)$, donde $X_0 = \{s, x_1, x_2, x_3\}$ y $X_1 = \{s, x_1\}$, son cortes mínimos de la red con una capacidad de paso igual a 4. Observemos, además, que en dicha red fluye el flujo máximo cuya magnitud es también igual a 4.

El siguiente teorema sobre el flujo máximo y el corte mínimo es el resultado principal referente al flujo estacionario máximo en la red.

Teorema 1. En cualquier red el valor máximo de un flujo de s a t es igual a la capacidad de paso mínima del corte que separa s de t .

Demostración. Del lema 1 se deduce que resulta suficiente comprobar la existencia del flujo f y del corte $(Y, \bar{Y})$, para los cuales el valor del flujo y la capacidad de paso del corte son iguales. En este caso el flujo será máximo y el corte, mínimo.

Tomemos el flujo máximo f (tal flujo existe, evidentemente, aunque pueden haber varios de ellos) y definamos de una manera recurrente algún corte:

a) $s \in Y$;

b) Si $x \in Y$, y $f(x, y) < c(x, y)$, entonces $y \in Y$; si $x \in Y$, y $f(y, x) > 0$, entonces $y \in Y$.

Demostremos que el conjunto obtenido $(Y, \bar{Y})$ es un corte en la red $G = (X, U)$. Para ello basta demostrar que $t \in \bar{Y}$.

Supongamos que esto no es así. Entonces, de la definición del conjunto

Y se desprende que existe un camino de s a t ($s, x_1, x_2, \dots, x_n, t$) que posee la propiedad de que para todos los arcos (x_i, x_{i+1}) del camino citado tiene lugar o bien

$$f(x_i, x_{i+1}) < c(x_i, x_{i+1}), \quad (3)$$

o bien

$$f(x_{i+1}, x_i) > 0. \quad (4)$$

Sea ε_1 el mínimo de la diferencia $c - f$, calculada en aquellos arcos del camino $(s, x_1, \dots, x_n, t)$, para los cuales se verifica la desigualdad (3), y sea ε_2 el mínimo de f en todos los arcos inversos, en los que se verifica la desigualdad (4). Pongamos $\varepsilon = \min\{\varepsilon_1, \varepsilon_2\} > 0$. Modifiquemos ahora el flujo f del modo siguiente: aumentemos f en la magnitud ε en los arcos del camino $(s, x_1, \dots, x_n, t)$, sobre los cuales se verifican las desigualdades (3) y disminuyamos f en la magnitud ε en los arcos inversos, sobre los cuales se verifica (4). Una función, definida de este modo, será, evidentemente, el flujo en la red de s a t , cuyo valor es igual a $v + \varepsilon$. Mas, esto contradice la suposición de que el flujo f es máximo. Por consiguiente, $t \in \overline{Y}$, y el conjunto $(Y, \overline{Y})$ es el corte que separa s de t .

De la definición de Y se deduce:

$$\begin{aligned} f(x, \bar{x}) &= c(x, \bar{x}) \text{ para } (x, \bar{x}) \in (Y, \overline{Y}); \\ f(\bar{x}, x) &= 0 \text{ para } (\bar{x}, x) \in (\overline{Y}, Y). \end{aligned}$$

Por consiguiente,

$$v = f(Y, Y) - f(\overline{Y}, Y) = c(Y, Y).$$

El teorema está demostrado.

Del lema demostrado 1 y del teorema 4 pueden deducirse algunos corolarios.

Diremos que cierto itinerario de s a t aumenta el flujo f , si para cada arco U o bien $f(U) < c(U)$, o bien $f > 0$ en el arco inverso de U .

Corolario 1. El flujo f es máximo, cuando y sólo cuando no existe ningún itinerario que aumente el flujo f .

El arco (x, y) se llamará *saturado de flujo f* , si $f(x, y) = c(x, y)$, y *libre del flujo*, si $f(x, y) = 0$.

Corolario 2. Un corte $(Y, \overline{Y})$ es mínimo, si y sólo si cada flujo máximo f satura todos los arcos del corte $(Y, \overline{Y})$ y deja libres todos los arcos pertenecientes a $(\overline{Y}, Y)$.

El siguiente teorema demuestra que el corte mínimo $(Y, \overline{Y})$ construido en el teorema 4 no depende, en realidad, de cómo se elige el flujo máximo f .

Teorema 5. Supongamos que $(Z, \overline{Z})$ es un corte mínimo arbitrario; f , cierto flujo máximo y $(Y, \overline{Y})$, un corte construido por el flujo f en el teorema 4. Entonces

$$Y \subseteq Z.$$

Demostración. Sea $Y \not\subseteq Z$; entonces $Y \in Z \subset Y$. Demostremos la siguiente afirmación: $(Y \cap Z, \overline{Y \cap Z})$ es también un corte mínimo. Para la demostración hagamos uso del corolario 2. Sea $x \in Y \cap Z$, y $\bar{x} \in \overline{Y \cap Z}$. Por consiguiente, $x \in Y$; $x \in Z$; $\bar{x} \in \overline{Z \cup Y}$, y, por tanto, $(x, \bar{x})$ pertenece por lo menos a uno de los cortes: o bien a $(Y, \overline{Y})$, o bien a $(Z, \overline{Z})$ y entonces, de acuerdo con el corolario 2, el flujo máximo satura el arco $(x, \bar{x})$. Análogamente se puede demostrar que el flujo máximo arbitrario f deja libre el arco $(\bar{x}, x)$. Por consiguiente, $(Y \cap Z, \overline{Y \cap Z})$ es un corte mínimo.

Supongamos que el vértice $x \in Y$, y $x \notin Y \cap Z$. Entonces, $x \neq s$, y existe un camino de s a x , digamos, $(s = x_1, x_2, \dots, x_n = x)$ tal que cada uno de sus arcos o bien está saturado de flujo f , o bien el arco inverso no es libre del flujo. Por cuanto $s \in Y \cap Z$, y $x \in \overline{Y \cap Z}$, se encontrará un par (x_i, x_{i+1}) , en el que $x_i \in Y \cap Z$ y $x_{i+1} \in \overline{Y \cap Z}$. Demostrado el hecho de que $(Y \cap Z, \overline{Y \cap Z})$ es el corte mínimo, podemos concluir, de acuerdo con el corolario 2, que todos los arcos rectos de él están saturados de flujo f , y todos los arcos inversos de él son libres del flujo f . Hemos llegado, pues, a una contradicción y, por lo tanto, $Y \subseteq Z$. El teorema está demostrado.

Varias entradas y salidas. Supongamos que el conjunto de vértices x de la red (X, U) está partido en tres conjuntos:

S es un conjunto de entradas,

T , un conjunto de salidas,

R , un conjunto de vértices intermedios.

Por flujo de S a T se entiende una función numérica f que está definida en U y satisface las condiciones:

$$f(x, X) - f(X, x) = 0; \quad x \in R;$$

$$f(S, X) - f(X, S) = v;$$

$$f(T, X) - f(X, T) = -v;$$

$$0 \leq f(x, y) \leq c(x, y).$$

Ensanchemos la red (X, U) , agregando dos vértices s^* y t^* y todos los arcos (s^*, S) y (T, t^*) . Obtendremos una nueva red (X^*, U^*) . Definamos adicionalmente la función de la capacidad de paso c en U^* , poniendo

$$c^*(s^*, x) = \infty; \quad x \in S;$$

$$c^*(x, t^*) = \infty; \quad x \in T;$$

$$c^*(x, y) = c(x, y); \quad (x, y) \in U.$$

Es fácil ver que el estrechamiento f de cualquier flujo f^* de s^* a t^* en la red (X^*, U^*) representa el flujo de S a T en la red (X, Y) . Y viceversa, cualquier flujo f de S a T en la red (X, U) continúa como el flujo f^* de s^* a t^* en la red (X^*, U^*) :

$$f^*(s^*, x) = f(x, X) - f(X, x); \quad x \in S;$$

$$f^*(x, t^*) = f(X, x) - f(x, X); \quad x \in T;$$

$f^*(x, y) = f(x, y)$ en todos los demás casos.

Por consiguiente, el problema del flujo máximo de S a T en la red (X, U) es equivalente al del flujo máximo en una red ensanchada con una entrada y una salida.

Método de marcación para resolver problemas del flujo máximo. Existen varios algoritmos para hallar el flujo máximo. No obstante, en la base de la mayoría de ellos yace el llamado "método de marcación". Apoyándonos en la demostración del teorema 1 y de sus corolarios, damos a conocer este método eficaz de construcción del flujo máximo y del corte mínimo.

Para asegurar la terminación del proceso, convengámonos en considerar que la función de la capacidad de paso c toma sólo valores de números enteros. En la práctica esto no es una restricción sustancial, puesto que las capacidades de paso racionales pueden reducirse a los números enteros a cuenta del cambio de la escala.

Sea dado un flujo f de números enteros en la red (X, U) . (En el instante inicial el flujo f puede ser nulo). Asignemos a los vértices de la red dada unas marcas que pueden tener la forma (x^+, ε) ó (x^-, ε) , donde $x \in X$, y ε es un número natural, o bien ∞ . Al realizarse la operación A, pueden ser posibles tres estados de un vértice: a) el vértice no está marcado; b) el vértice está marcado, pero no observado; c) el vértice está marcado y observado.

Operación A (marcación). La entrada S recibe la marca $(-\infty)$, es decir, $\varepsilon(S) = \infty$. Ahora la entrada está marcada, pero no observada, mientras que todos los demás vértices no están marcados. En general, elijamos cualquier vértice x que está marcado, pero no observado. Supongamos que tiene la marca $(x^+, \varepsilon(x))$. Entonces, a todos los vértices y , que no están marcados y para los cuales $f(x, y) < c(x, y)$, les asignamos la marca $(x^+, \varepsilon(y))$, donde

$$\varepsilon(y) = \min\{\varepsilon(x), c(x, y) - f(x, y)\}.$$

Ahora los vértices y están marcados, pero no observados. A los vértices que después de esto quedan no marcados, pero para los cuales $f(x, y) > 0$, les asignamos la marca $(x^-, \varepsilon(y))$, donde

$$\varepsilon(y) = \min\{\varepsilon(x), f(x, y)\}.$$

Tales vértices y están ahora marcados y no observados, y el vértice x está marcado y observado. Esta operación se repite hasta que resulte marcada la salida t , o bien hasta que no sea posible marcar ningún vértice, mientras que la salida quede no marcada. En el último caso el proceso de construcción de f se considera terminado. En el primer caso pasamos a la operación B.

Operación B (cambio del flujo). Supongamos que la salida t tiene la marca $(y^+, \varepsilon(t))$; entonces $f(y, t)$ se sustituye por $f(y, t) + \varepsilon(t)$; en cambio, si t está marcado con $(y^-, \varepsilon(t))$, entonces $f(t, y)$ se sustituye por $f(t, y) - \varepsilon(t)$. Luego, en cualquiera de estos casos pasamos al vértice y . En

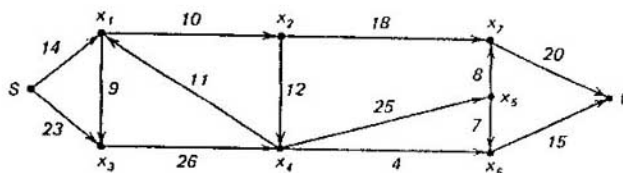


Fig. 6.24.

general, si y tiene la marca $(x^+, c(y))$, entonces $f(x, y)$ se sustituye por $f(x, y) + c(t)$, y si el vértice y está marcado con $(x^-, c(y))$, entonces $f(y, x)$ se sustituye por $f(y, x) - c(t)$, después de lo cual se pasa al vértice x . Conseguida la entrada s , el cambio del flujo se da por terminado. Se borran todas las marcas y de nuevo se pasa a la operación A para el nuevo aumento del flujo.

En el proceso de marcación, si pasamos de la operación A a la B, esto quiere decir que hemos encontrado el camino de s a t , que aumenta el flujo f . En cambio, si la operación A está terminada, mientras que la salida, no marcada, entonces el flujo fue máximo y el conjunto de arcos que unen los vértices marcados con los no marcados, forman el corte mínimo.

Veamos ahora un ejemplo que ilustra la aplicación del método de marcación para resolver los problemas de flujo máximo. Sea una red expuesta en la fig. 6.24 con la entrada s y la salida t . Las capacidades de paso de los arcos vienen indicadas también en dicha figura. Se requiere hallar el flujo máximo de s a t .

Operación A. Asignemos al vértice s una marca $(-\infty)$.

El conjunto de vértices $\{x_j | x_j \in \Gamma(s), f(s, x_j) < c(s, x_j), x_j \text{ no está marcado}\}$ es $\{x_1, x_3\}$, al vértice x_1 se le asigna la marca $(s^+, \min\{\infty, 14-0\})$, es decir, $(s^+, 14)$; al vértice x_3 se le asigna la marca $(s^+, \min\{\infty, 23-0\})$, es decir, $(s^+, 23)$.

El conjunto $\{x_j | x_j \in \Gamma^{-1}(s), f(x_j, s) > 0, x_j \text{ no están marcados}\}$ es vacío. Así pues, el vértice s está marcado y observado; los vértices x_1 y x_3 están marcados y no observados, y todos los demás vértices, no marcados.

Observemos ahora el vértice x_1 $\{x_j | x_j \in \Gamma(x_1), f(x_1, x_j) < c(x_1, x_j), x_j \text{ no está marcado}\} = \{x_2\}$, para x_2 de marca sirve $(x_1^+, \min\{14, 10-0\}) = (x_1^+, 10)$. $\{x_j | x_j \in \Gamma^{-1}(x_1), f(x_j, x_1) > 0, x_j \text{ no está marcada}\} = \emptyset$. Ahora los vértices s y x_1 están marcados y observados, mientras que los x_2 y x_3 , marcados, pero no observados.

Al tomar para la observación el vértice x_2 , obtenemos las siguientes marcas: $(x_2^+, 10)$ para x_4 y $((x_2^+, 10)$, para x_7 .

Al tomar para la observación x_3 , llegamos a que no se puede poner ninguna marca. Al continuar la observación desde el vértice x_4 , obtenemos las siguientes marcas: $((x_4^+, 10)$ para x_5 , $((x_4^+, 4)$ para x_6 y $(x_6^+, 4)$, para t .

Operación B. Aquí obtenemos: $x^- = t$; $f(x_6, t) = 0 + 4 = 4$; $x^- = x_6$;

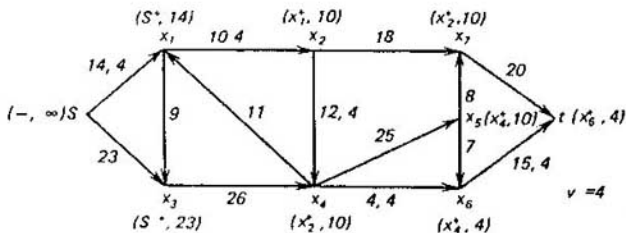


Fig. 6.25.

$f(x_4, x_6) = 0 + 4 = 4$; $x^- = x_4$; $f(x_2, x_4) = 0 + 4 = 4$; $x^- = x_2$; $f(x_1, x_2) = 0 + 4 = 4$, $x^- = x_1$;

La forma del flujo al final de la operación B y las marcas de los vértices antes de que ellas quedan borradas se muestran en la fig. 6.25, donde el primer número en el arco indica la capacidad de paso, y el segundo, el valor de la función f .

Borrando las marcas de los vértices y volviendo a la operación A para la segunda pasada, obtendremos nuevas marcas de los vértices (los vértices marcados, pero no observados, los observaremos en el orden de crecimiento de sus números ordinales).

Operación A. La marca para s será $(-, \infty)$; para x_1 , $(s^+, 10)$, y para x_3 , $(s^+, 23)$; ahora el vértice s está marcado y observado. La marca para x_2 será $(x_1^+, 6)$ y el vértice x_1 queda marcado y observado. La marca para x_4 será $(x_2^+, 6)$ y para x_7 , $(x_2^+, 6)$. Ahora el vértice x_2 está marcado y observado, igual que el vértice x_3 . La marca para x_5 será $(x_4^+, 6)$ y el vértice x_4 se hace marcado y observado. La marca para x_6 será $(x_5^+, 6)$, después de lo cual el vértice x_5 será marcado y observado. Por último, la marca para t será $(x_6^+, 6)$.

Operación B. El nuevo flujo ha aumentado del modo siguiente:

$f(x_6, t) = 4 + 6 = 10$; $f(x_5, x_6) = 0 + 6 = 6$; $f(x_4, x_5) = 0 + 6 = 6$;
 $f(x_2, x_4) = 4 + 6 = 10$; $f(x_1, x_2) = 4 + 6 = 10$; $f(s, x_1) = 4 + 6 = 10$.

Todos los demás valores de la función f quedan los mismos. La forma nueva del flujo y las marcas de los vértices se muestran en la fig. 6.26.

Continuando el proceso, obtenemos tras cada paso del algoritmo, los flujos y las marcas que se muestran sucesivamente en las figs. 6.27—6.29.

El algoritmo finaliza su trabajo cuando ningún vértice puede ser marcado, como tampoco la salida t ; las marcas "finales" se muestran en la fig. 6.30. El flujo expuesto en la fig. 6.30 es, por eso, el máximo del valor 29, mientras que el correspondiente corte mínimo en este dibujo está representado con una línea punteada.

Ejercicios.

1. Hállese, empleando el método de marcación, el flujo máximo de la entrada s a la salida t en las redes expuestas en las figs. 6.31 a y 6.31b. Junto a los arcos en la fig. 6.31 se indican las capacidades de paso y los valores corrientes de la función del flujo f .

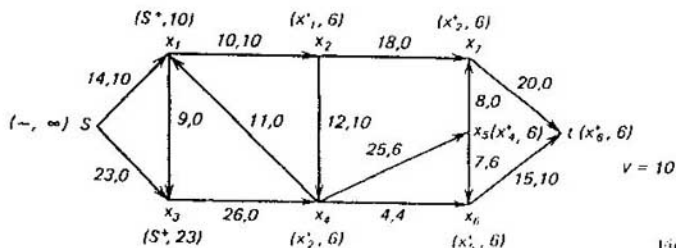


Fig. 6.26.

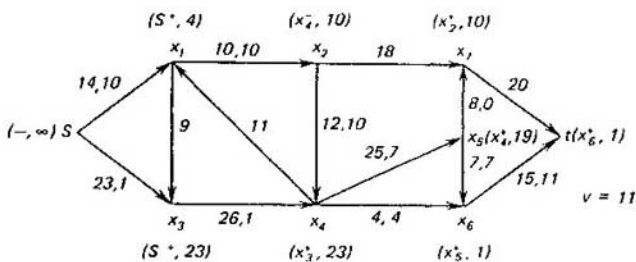


Fig. 6.27.

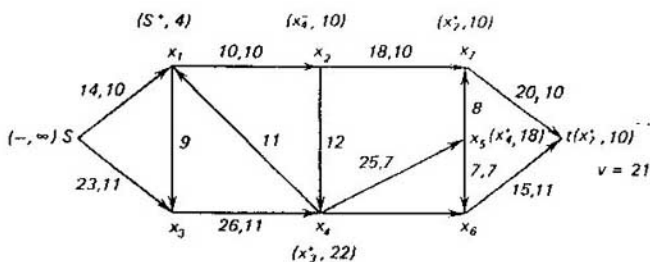


Fig. 6.28.

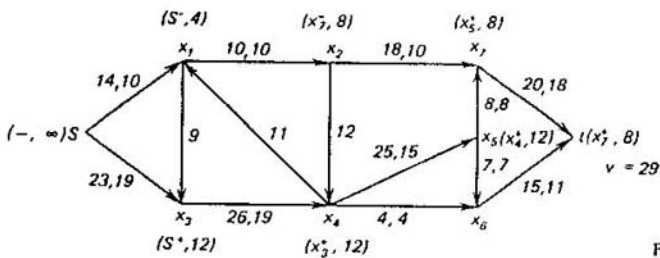


Fig. 6.29.

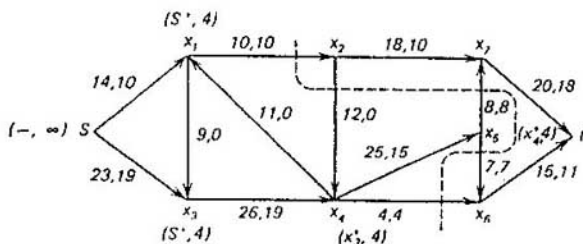


Fig. 6.30.

2. Hállese el corte mínimo que separa la entrada s de la salida t para las redes del ejercicio 1. ¿Saturará el flujo máximo todos los arcos de este corte?

3. Enúnciense las condiciones, bajo las cuales en una red arbitraria existe un camino, que aumenta el flujo, de la entrada s a la salida t , que sea compuesto íntegramente de arcos inversos (arcos rectos).

Del algoritmo de marcación se deduce que si el flujo original era de números enteros (es decir, $f(u)$ representaba números enteros para cualquier $u \in U$), entonces todos los flujos que siguen después de los aumentos resultarán ser de números enteros. Por consiguiente, es lícito el

Teorema 6. Si la función de la capacidad de paso c es de números enteros, existe un flujo máximo que también es de números enteros.

Restricciones inferiores de los flujos de arco. En la definición de flujo, en lugar de la desigualdad (2) podemos hacer

$$l(x, y) \leq f(x, y) \leq c(x, y), \quad (5)$$

donde $l(x, y)$ es una función real definida sobre los arcos del conjunto U , y $0 \leq l(x, y) \leq c(x, y)$.

El proceso de marcación se extiende inmediatamente a este caso con algunas variaciones insignificantes. Siendo vigentes las restricciones (5), se demuestra con facilidad el análogo del teorema básico 4.

Teorema 7. Si existe una función f que satisface, para cierto v , las desigualdades (5) y ecuaciones (1), entonces el valor máximo del flujo de s a t es igual al mínimo de la diferencia $c(\bar{Y}, \bar{Y}) - l(\bar{Y}, \bar{Y})$ tomado en todos los cortes $(Y, \bar{Y})$ de la red.

Más abajo se demostrarán los teoremas que proporcionan las condi-

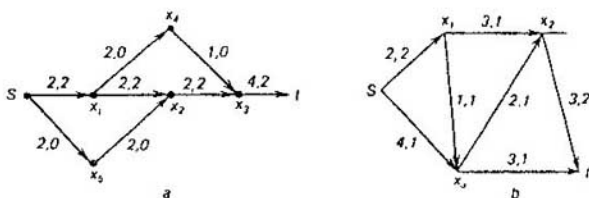


Fig. 6.31.

ciones necesarias y suficientes para la existencia de los flujos en las redes, las cuales satisfacen diferentes desigualdades lineales.

Teorema de demanda y oferta. Sea $G = (X, U)$ una red arbitraria con la capacidad de paso $c = c(x, y)$, donde $(x, y) \in U$, y X está partido en tres conjuntos: $X = S \cup R \cup T$. S son las entradas, R son los vértices intermedios, T son las salidas. A todo vértice $x \in S$ se le asigna un número $a(x)$, y a cada vértice $y \in T$ se le asigna el número $b(y)$; $a(x) \geq 0$, $x \in S$; $b(y) \geq 0$, $y \in T$. El número $a(x)$ puede entenderse como oferta de la mercancía en la entrada x ; $b(y)$, como la demanda de la mercancía en la salida y . Surge el siguiente problema: ¿cuáles son las condiciones necesarias y suficientes, bajo las cuales la demanda en las salidas puede ser satisfecha por la oferta en las entradas? En el lenguaje algebraico esto puede expresarse así:

$$\begin{aligned} f(x, X) - f(X, x) &\leq a(x); \quad x \in S; \\ f(x, X) - f(X, x) &= 0; \quad x \in R; \\ f(X, x) - f(x, X) &\geq b(x); \quad x \in T; \\ 0 &\leq f(x, y) \leq c(x, y). \end{aligned} \quad (6)$$

Teorema 8. Las restricciones (6), donde $a(x) \geq 0$ y $b(x) \geq 0$ son admisibles cuando y sólo cuando para cada subconjunto $Y \subseteq X$:

$$b(T \cap \bar{Y}) - a(S \cap \bar{Y}) \leq c(Y, \bar{Y}). \quad (7)$$

Demostración. Demostremos al principio la necesidad de la desigualdad (7). Supongamos que existe un flujo f que satisface las condiciones (6). Al sumar estas desigualdades respecto de todos los vértices $x \in \bar{Y}$, obtenemos una desigualdad

$$b(T \cap \bar{Y}) - a(S \cap \bar{Y}) \leq f(X, \bar{Y}) - f(\bar{Y}, X). \quad (8)$$

$X = Y \cup \bar{Y}$, por lo cual (8) puede ser transformada:

$$b(T \cap \bar{Y}) - a(S \cap \bar{Y}) \leq f(Y, \bar{Y}) - f(\bar{Y}, Y).$$

Haciendo uso del resultado del lema 1, demostramos la necesidad de la condición 7.

Demostremos ahora la suficiencia de la condición (7). Hagamos ensanchar la red $G = (X, U)$, uniendo la entrada ficticia s , la salida t y los arcos (s, S) y (T, t) . Obtendremos una red nueva X^*, U^* . La función que caracteriza la capacidad de paso sobre U^* la definimos según las fórmulas:

$$\begin{aligned} c^*(s, x) &= a(x); \quad x \in S; \\ c^*(x, t) &= b(x); \quad x \in T; \\ c^*(x, y) &= c(x, y); \quad (x, y) \in U. \end{aligned}$$

La validez de la desigualdad (7) significa que el corte (T, t) en la red (X^*, U^*) es mínimo. Demostremoslo. Sea $(Y^*, \bar{Y}^*)$ un corte arbitrario que separa

s de t . Pongamos $Y = Y^* \setminus s$, $\bar{X} = X^* \setminus t$, entonces

$$\begin{aligned} c^*(Y^*, \bar{Y}) - c^*(T, t) &= c^*(Y, t) + c^*(s, \bar{Y}) + \\ &+ c^*(Y, \bar{Y}) - c^*(T, t) = b(T \cap Y) + a(S \cap \bar{Y}) + c(Y, \bar{Y}) - b(T) = \\ &= -b(T \cap \bar{Y}) + a(S \cap \bar{Y}) + c(Y, \bar{Y}). \end{aligned}$$

Aplicando el teorema de flujo máximo y corte mínimo, llegamos a que el cumplimiento de la condición (7) tiene por resultado la existencia del flujo f^* de s a t en la red (X^*, U^*) que satura todos los arcos del corte (T, t) . En tal caso el flujo f que representa un estrechamiento del flujo f^* en la red (X, U) satisface las condiciones (6), puesto que

$$\begin{aligned} a(x) &\geq f^*(s, x) = f^*(x, X) - f^*(X, x) = f(x, X) - f(X, x); \quad x \in S; \\ b(x) &= f^*(x, t) = f^*(X, x) \text{ y } f^*(x, X) = f(X, x) - f(x, X); \quad x \in T. \end{aligned}$$

El teorema está demostrado.

Teorema de circulación. Demostremos un teorema más, en el cual se trata de la existencia de las circulaciones, es decir, de los flujos que están privados de entradas y salidas y que están limitados en los arcos por las fronteras superiores e inferiores prefijadas.

Supongamos que para la red (X, U) están prefijadas l y c , es decir, las funciones de frontera inferior y superior, respectivamente, en U . La circulación admisible en la red (X, Y) es f , que está definida sobre U y que satisface las condiciones:

$$f(x, X) - f(X, x) = 0; \quad x \in X; \quad (9)$$

$$0 \leq l(x, y) \leq f(x, y) \leq c(x, y); \quad (x, y) \in U. \quad (10)$$

Teorema 9. Para que las restricciones (9) y (10) sean admisibles, es necesario y suficiente que para cualquier $Y \subseteq X$ se cumpla la desigualdad

$$c(Y, \bar{Y}) \geq l(\bar{Y}, Y).$$

Demostración. Hagamos ensanchar la red (X, U) añadiendo dos vértices s y t , y un conjunto de arcos (s, X) y (X, t) . En la nueva red (X^*, U^*) definamos una función que tiene una capacidad de paso según las fórmulas:

$$c^*(x, y) = c(x, y) - l(x, y); \quad (x, y) \in U;$$

$$c^*(s, x) = l(X, x); \quad x \in X;$$

$$c^*(x, t) = l(x, X); \quad x \in X.$$

Si en la red (X, U) existe una circulación admisible f , entonces ésta engendra el flujo f^* de s a t en la red (X^*, U^*) :

$$f^*(x, y) = f(x, y) - l(x, y); \quad (x, y) \in U;$$

$$f^*(s, x) = l(X, x); \quad x \in X;$$

$$f^*(x, t) = l(x, X); \quad x \in X.$$

Al contrario, si en la red (X^*, U^*) existe un flujo $f^*(x, y)$, entonces en la red (X, U) puede determinarse la circulación $f(x, y)$ al poner:

$$f(x, y) = f^*(x, y) + l(x, y) \text{ para } (x, y) \in U.$$

Se comprueba con facilidad el cumplimiento de las relaciones (9) y (10).

Así pues, el problema se ha reducido a la cuestión sobre la existencia del flujo de s a t en la red (X^*, U^*) de valor $l(X, X)$. La condición necesaria y suficiente para la existencia del flujo citado de s a t de valor $l(X, X)$ consiste en que las capacidades de paso de todos los cortes sean no menos de $l(X, X)$.

Sea $(Y, \bar{Y}^*)$ un corte que separa s de t en la red (X^*, U^*) . Definamos el conjunto $Y \subseteq X$ del modo siguiente:

$$Y = Y^* \setminus s; \quad \bar{Y} = \bar{Y}^* \setminus t.$$

Entonces:

$$\begin{aligned} c^*(Y^*, \bar{Y}^*) &= c^*(Y \cup s, \bar{Y} \cup t) = c^*(Y, \bar{Y}) + c^*(s, \bar{Y}) + c^*(Y, t) = \\ &= c(Y, \bar{Y}) - l(Y, \bar{Y}) + l(X, \bar{Y}) + l(Y, X) = c(Y, \bar{Y}) + \\ &+ l(\bar{Y}, \bar{Y}) + l(Y, X) = c(Y, \bar{Y}) + l(X, X) - l(\bar{Y}, Y). \end{aligned}$$

Por consiguiente, $c^*(Y^*, \bar{Y}^*) \geq l(X, X)$ cuando y sólo cuando

$$c(Y, \bar{Y}) \geq l(\bar{Y}, Y).$$

Con ello queda demostrado el teorema.

Por último demostremos el análogo del teorema de König (véase § 4.1), ilustrando de este modo su conexión con el teorema sobre el flujo máximo y el corte mínimo.

Sea $G = G(S \cup T, U)$ un grafo bipartido arbitrario sobre un conjunto de vértices $S \cup T$ con un conjunto de arcos U , cuyos arcos están todos orientados de S en T . Se denomina combinación de pares en el grafo G un subconjunto E de aristas, de las cuales ningún par tiene vértices comunes. Un conjunto de vértices se denomina (S, T) -separador, si la eliminación en el grafo de los citados vértices junto con los arcos incidentes con ellos rompe todos los caminos que van desde los vértices del conjunto S a los vértices del conjunto T .

Teorema 10. Sea $G = G(S \cup T, U)$ un grafo bipartido. Entonces, el número máximo de arcos de la combinación de pares en el grafo G es igual al número mínimo de vértices en el conjunto (S, T) -separador.

Demostración. Juntemos al grafo en consideración G dos vértices s y t , y también los arcos (s, S) y (T, t) , y para todos los $x \in S$ e $y \in T$ definamos las capacidades de paso de los arcos del grafo:

$$\begin{aligned} c(s, x) &= 1; \quad x \in S; \\ c(x, t) &= 1; \quad x \in T; \\ c(x, y) &= \infty; \quad (x, y) \in U. \end{aligned} \tag{11}$$

Sea f un flujo máximo arbitrario de s a t , que se mide en números enteros, y sea $(Y, \bar{Y})$ cierto corte mínimo que separa s de t . Teniendo presente (11), concluimos que los arcos del conjunto $I = \{(x, y) \in U; f(x, y) = 1\}$ no tienen vértices comunes de dos en dos. El conjunto $D = (S \cap \bar{Y}) \cup (T \cap Y)$ es (S, T) -separador, con la particularidad de que D se encuentra en una correspondencia biunívoca con los arcos del corte mínimo $(Y, \bar{Y})$. Al hacer uso del teorema sobre el flujo máximo y el corte mínimo, llegamos a que si v es el valor del flujo máximo f , entonces $|I| = |D| = v$. Por otra parte, el número máximo de arcos del grafo G que no tienen vértices comunes de dos en dos no es, obviamente, superior al número mínimo de vértices en cualquier conjunto (S, T) -separador. El teorema está demostrado.

A las relaciones existentes entre el teorema sobre el flujo máximo y el corte mínimo y otros problemas combinatorios extremales volveremos una vez más en el § 8.1.

El dominio de los problemas extremales de carácter discreto es enorme. Los esfuerzos de muchos matemáticos están dirigidos a resolverlos. Han sido elaborados algoritmos para la resolución de algunas clases de problemas. No obstante, no existe todavía una teoría íntegra y única en dicho dominio.

CAPÍTULO 7 MÉTODOS PROBABILÍSTICOS EN EL ANÁLISIS COMBINATORIO

7.1. EJEMPLOS DE APLICACIÓN DE LOS MÉTODOS PROBABILÍSTICOS

En el § 2.4 se han estudiado las aplicaciones del aparato combinatorio de funciones generatrices a los problemas de la teoría de las probabilidades (para un conocimiento más detallado de esta última se recomienda el libro [82]). En el presente capítulo el planteamiento del problema será inverso. Daremos a conocer un método (o, incluso, un grupo de métodos) de resolución de los problemas combinatorios por medios de carácter probabilístico. La esencia del método consiste en que los objetos combinatorios que se investigan se interpretan como sucesos en cierto espacio probabilístico discreto. Se demuestra que en este espacio la probabilidad de que un objeto, poseedor de las propiedades dadas, se realice es positiva. De este modo se demuestra la existencia de tal objeto, aunque dicho objeto no se construye explícitamente. El método no es perfecto por su carácter no constructivo, sin embargo, en muchos problemas combinatorios resulta suficiente establecer el propio hecho de existencia de un objeto. El método que se describe emplea el aparato desarrollado de la teoría de las probabilidades, lo que permite simplificar razonamientos engorrosos.

Explicaremos la esencia del método con ejemplos concretos.

Estimaciones de los números de Ramsey. Enunciemos el teorema de Ramsey del párrafo 3.3 para un caso particular en que $r = t = 2$, $q_1 = q_2 = k$.

Teorema 1. Existe un número natural mínimo $n = R(k)$ tal que para cualquier coloración en dos colores de las aristas de un grafo no orientado completo de n vértices se encontrará un subgrafo completo de k vértices, cuyas aristas están coloreadas de un mismo color.

Demostremos por el método probabilístico la estimación inferior para los números de Ramsey $R(k)$.

Teorema 2.

$$R(k) \geq k \cdot 2^{k/2} \cdot \left(\frac{1}{\sqrt{2e}} + o(1) \right) \text{ para } k \rightarrow \infty.$$

Demostración. Analicemos un grafo completo con un conjunto de vértices X , donde $|X| = n$. A título de espacio probabilístico elijamos el conjunto de todas las coloraciones de sus aristas en 2 colores: rojo y azul. Todas las coloraciones son equiprobables, es decir, cada arista se colorea, independientemente de las otras, de color rojo con una probabilidad igual a $1/2$, y de color azul, con una probabilidad de $1/2$. Examinemos un k -subconjunto $Y \subseteq X$. La probabilidad de que todas las aristas del subgrafo

generado por Y estén coloreadas de rojo será igual a $2^{-\binom{k}{2}}$. Para el color azul esta probabilidad es la misma. Por eso, la probabilidad de que el subconjunto dado de los vértices Y genere un subgrafo monocromático es igual a $2 \cdot 2^{-\binom{k}{2}} = 2^{1-\binom{k}{2}}$. El conjunto Y puede ser elegido por $\binom{n}{k}$ métodos. Por eso, la probabilidad de la existencia del k -subconjunto Y , que genera un subgrafo monocromático no es superior a $\binom{n}{k} \cdot 2^{1-\binom{k}{2}}$. Por consiguiente, si los números n y k son de tal género que $\binom{n}{k} \cdot 2^{1-\binom{k}{2}} < 1$, entonces el suceso se realiza con la probabilidad positiva $p = 1 - \binom{n}{k} \cdot 2^{1-\binom{k}{2}}$, cuando ningún subgrafo generado de k vértices es monocromático. En virtud de la finitud del conjunto de coloraciones esto quiere decir que existe una coloración, para la cual no existen subgrafos completos monocromáticos de k vértices. Así pues, de $\binom{n}{k} < 2^{\binom{k}{2}-1}$ se deduce: $n < R(k)$. Por consiguiente,

$$\binom{R(k)}{k} \geq 2^{\binom{k}{2}-1}.$$

Haciendo uso del desarrollo de Stirling $k! = k^k e^{-k} \cdot \sqrt{2\pi k} \cdot (1 + o(1))$ para $k \rightarrow \infty$, obtendremos

$$R(k)^k \geq k! \binom{R(k)}{k} \geq k! \cdot 2^{\binom{k}{2}-1}.$$

$$R(k) \geq \left(k^k e^{-k} \cdot \sqrt{2\pi k} \cdot 2^{\frac{k(k-1)}{2}-1} \right)^{\frac{1}{k}} \cdot (1 + o(1)) = \frac{k \cdot 2^{k/2}}{\sqrt{2e}} \cdot (1 + o(1)),$$

lo que se requería demostrar.

Propiedad $\mathcal{B}$ (véase la definición en el § 1.5). Designemos con $m(k)$ el número mínimo de miembros en el sistema de k -subconjuntos que no poseen la propiedad $\mathcal{B}$. Es fácil ver que el sistema de todos los k -subconjuntos del $(2k-1)$ -conjunto no posee la propiedad $\mathcal{A}$, por lo cual

$$m(k) \leq \binom{2k-1}{k}.$$

La estimación de debajo para $m(k)$ la obtendremos por el método de los primeros momentos.

Teorema 3.

$$m(k) \geq 2^{k-1}.$$

Demostración. Veamos un conjunto V y un sistema $A = \{S_1, S_2, \dots, S_m\}$ de sus k -subconjuntos que no posee la propiedad $\mathcal{A}$, donde $m = m(k)$.

Sea $C = (V_1, V_2)$ una partición del conjunto V , es decir, $V = V_1 \cup V_2$, $V_1 \cap V_2 = \emptyset$. Pongamos

$$f_{ij}(C) = \begin{cases} 1, & \text{si } S_j \subseteq V_i, \\ 0, & \text{si } S_j \not\subseteq V_i, \end{cases}$$

donde $i = 1, 2$; $j = 1, 2, \dots, m$. Entonces

$$\beta(C) = \sum_{j=1}^m \sum_{i=1}^2 f_{ij}(C)$$

es el número de miembros del sistema A que están íntegramente contenidos en una de las partes de la partición C .

Sea ahora ζ una partición accidental, donde cada elemento del conjunto V pertenece con la probabilidad $1/2$, independientemente de los otros elementos, a la primera parte de la partición, y con la misma probabilidad $1/2$, a la segunda parte. Entonces, cuando $i = 1, 2$ para cualquier subconjunto S_j de k elementos, tenemos:

$$Mf_{ij}(C) = P\{f_{ij}(C) = 1\} = 2^{-k}.$$

Calculemos la esperanza matemática $\beta(C)$:

$$M\beta(C) = \sum_{j=1}^m \sum_{i=1}^2 Mf_{ij}(C) = \sum_{j=1}^m \sum_{i=1}^2 2^{-k} = m \cdot 2^{1-k}.$$

Por cuanto A no posee la propiedad $\mathcal{B}$, entonces $\beta(C) \geq 1$ para cualquier partición C y, por lo tanto,

$$M\beta(C) \geq 1.$$

Así pues, $m \cdot 2^{1-k} \geq 1$, lo que se trataba de demostrar.

Problema sobre la cantidad máxima de números cuyas sumas son todas diferentes. Designemos con $g(n)$ la cantidad máxima k de números naturales $a_1, \dots, a_k$, no superiores a n , cuyas sumas $\sum_{i \in S} a_i$ son todas diferentes para diferentes $S \subseteq \{1, \dots, k\}$. Está claro que los números $1, 2, 4, \dots, 2^{k-1}$ satisfacen esta condición cuando $k \leq 1 + \log_2 n$, por lo cual

$$g(n) \geq 1 + [\log_2 n].$$

Por otra parte, cualquiera de las sumas no sobrepasa kn . Si $k = g(n)$, se tienen en total $2^{k(n)}$ sumas diferentes y todas ellas están encerradas entre 0 y $ng(n)$, de donde $2^{g(n)} \leq ng(n)$. Por consiguiente

$$\begin{aligned} g(n) &\leq \log_2 n + \log_2 g(n) \leq \log_2 n + \log_2(\log_2 n + \log_2 g(n)) = \\ &= \log_2 n + \log_2(\log_2 n) + o(1). \end{aligned}$$

Esta estimación superior puede ser mejorada, si hacemos uso del método teórico-probabilístico de los segundos momentos.

Teorema 4.

$$g(n) \leq \log_2 n + \frac{1}{2} \log_2(\log_2 n) + o(1) \text{ para } n \rightarrow \infty.$$

Demostración. Supongamos que los números $a_1, \dots, a_k$ satisfacen las condiciones del problema. Examinemos una magnitud aleatoria $\xi = \sum_{i \in S} a_i$, donde S es un subconjunto elegido al azar del conjunto $\{1, \dots, k\}$. Cada uno de los elementos de dicho conjunto integra S independientemente de los otros elementos con una probabilidad igual a $1/2$. Por eso,

$$M\xi = \sum_{i=1}^k P\{i \in S\} \cdot a_i = \frac{1}{2} \sum_{i=1}^k a_i.$$

Cada par de elementos $\{i, j\}$, donde $1 \leq i < j \leq k$, integra el subconjunto S con la probabilidad igual a $1/4$. Por eso,

$$\begin{aligned} M\xi^2 &= M\left(\sum_{i \in S} a_i\right)^2 = \sum_{i=1}^k P\{i \in S\} \cdot a_i^2 + \sum_{1 \leq i < j \leq k} P\{i, j \in S\} \cdot 2a_i a_j = \\ &= \frac{1}{2} \sum_{i=1}^k a_i^2 + \frac{1}{4} \sum_{1 \leq i < j \leq k} 2a_i a_j = \frac{1}{4} \sum_{i=1}^k a_i^2 + \frac{1}{4} \left(\sum_{i=1}^k a_i\right)^2. \\ D\xi &= M\xi^2 - (M\xi)^2 = \frac{1}{4} \sum_{i=1}^k a_i^2. \end{aligned}$$

Emplearemos la desigualdad de Chébishev ($\sigma > 0$):

$$P\{|\xi - M\xi| < \sigma\} \geq 1 - \frac{D\xi}{\sigma^2} = 1 - \frac{1}{4\sigma^2} \sum_{i=1}^k a_i^2 \geq 1 - \frac{kn^2}{4\sigma^2}.$$

Por cuanto todas las sumas $\sum_{i \in S} a_i$ son distintas, el número de subconjuntos $S \subseteq \{1, \dots, k\}$, que satisfacen la condición $\left|\sum_{i \in S} a_i - M\xi\right| < \sigma$, no sobrepasa 2σ . Por consiguiente,

$$P\{|\xi - M\xi| < \sigma\} \leq 2\sigma \cdot \left(\frac{1}{2}\right)^k$$

y

$$1 - \frac{kn^2}{4\sigma^2} \leq \sigma \cdot 2^{1-k}$$

Al poner $\sigma = n\sqrt{k}$, obtenemos que $\frac{3}{8} \cdot 2^k \leq n\sqrt{k}$, de donde encontramos:

$$k \leq \log_2 n + \frac{1}{2} \log_2 (\log_2 n) + o(1) \text{ para } n \rightarrow \infty,$$

lo que se requería demostrar.

7.2. PROBLEMAS DE PLANIFICACIÓN DEL EXPERIMENTO

Examinemos el siguiente modelo matemático del problema de experimentos de rechazo. En un conjunto $X = \{x_1, \dots, x_n\}$, compuesto de n elementos, hay que hallar s elementos fijos, desconocidos de antemano (llamémoslos *defectuosos*). Concretamente estos pueden ser los problemas siguientes: búsqueda de las monedas falsas, determinación de los defectos en un televisor, comprobación de un cable telefónico por medio de un tester, localización de un hombre enfermo en un numeroso grupo de gente con ayuda del análisis de la sangre, etc.

Los experimentos se llevan a cabo del modo siguiente. Se toma cierto subconjunto $Y \subseteq X$. Como resultado de las pruebas (test) obtenemos una información que se interpreta según el planteamiento del problema. Por ejemplo, podemos determinar si pertenece al conjunto Y por lo menos uno de los s elementos defectuosos o bien cuántos, precisamente, elementos defectuosos se contienen en Y .

El juego de conjuntos Y_i ($i = 1, 2, \dots, N$) se llama *plan de experimentos*. Partiendo de los resultados de N experimentos se necesita hallar unívocamente todos los s elementos defectuosos, con la particularidad de que en la elección de $Y_i \subseteq X$ ($i = 1, 2, \dots, N$) pueden imponerse algunas restricciones.

Resulta natural dividir el problema de planificación de los experimentos en dos clases. A la primera clase se refiere la planificación *estática*, donde la elección del i -ésimo conjunto Y_i no depende de los resultados de los $i - 1$ experimentos anteriores. La segunda clase es la planificación *sucesiva*, donde los resultados de las comprobaciones anteriores influyen en el planteamiento del experimento de turno. Veamos primeramente el problema de búsqueda de un elemento defectuoso.

Sea $s = 1$, es decir, entre los elementos del conjunto $X = \{x_1, \dots, x_n\}$ hay exactamente un elemento defectuoso. Supongamos, además, que está prefijado el plan estático de los experimentos, es decir, el juego de conjuntos $Y_i \subseteq X$, $i = 1, \dots, N$. Basándose en dicho juego se construye la matriz $A = \|a_{ij}\|$ de dimensión $N \times n$:

$$a_{ij} = \begin{cases} 1, & \text{si } x_j \in Y_i, \\ 0, & \text{si } x_j \notin Y_i \end{cases} \quad (1 \leq j \leq n; 1 \leq i \leq N).$$

A la inversa, a base de esta matriz se restablecen unívocamente los conjuntos Y_i ($i = 1, 2, \dots, N$).

Lema 1. Sea $s = 1$. Los experimentos contruidos con arreglo a la matriz $A = \|a_{ij}\|$ restablecen unívocamente el elemento defectuoso cuando y sólo cuando todas las columnas de la matriz A son diferentes.

Demostración. Supongamos que en la matriz A la k -ésima columna es igual a la columna l -ésima. En este caso no podremos distinguir cuál de los elementos es defectuoso: x_k ó x_l , puesto que dichos elementos o bien

pertenecen simultáneamente al conjunto Y_m ($1 \leq m \leq N$), o bien no lo pertenecen, razón por la cual todas las columnas de la matriz A han de ser diferentes.

Al contrario, sea e_i el resultado de la i -ésima comprobación, es decir, $e_i = 1$, si el elemento defectuoso pertenece al conjunto Y_i del plan de experimentos construido a base de la matriz $A = [a_{ij}]$, y $e_i = 0$, en el caso contrario. Así pues, realizadas N comprobaciones, obtendremos un vector de resultados $\vec{e} = (e_1, e_2, \dots, e_N)^T$ (T es una operación de transposición). Si el elemento x_k es defectuoso, entonces $\vec{e}$ coincide con la k -ésima columna de la matriz $A = [a_{ij}]$. Por consiguiente, si todas las columnas de la matriz A son distintas, entonces partiendo del vector de los resultados restableceremos unívocamente el elemento defectuoso. La demostración está terminada.

Así pues, el problema de planificación estática consiste en hallar la matriz A , compuesta de ceros y unidades, tal que todas las n columnas cuyas sean distintas dos a dos y el número de filas, mínimo. Este número se denotará con $N_{\text{exp}}^{\text{est}}$.

Teorema 5. En el caso de un elemento defectuoso

$$N_{\text{exp}}^{\text{est}} = \lceil \log_2 n \rceil.$$

Demostración. Si todas las n columnas de una $(0, 1)$ -matriz A , compuesta de N filas, son distintas dos a dos, entonces su cantidad no sobrepasa la cantidad de todas las palabras binarias de longitud N , es decir, $n \leq 2^N$. Teniendo presente que N es un número entero y suponiendo $N = N_{\text{exp}}^{\text{est}}$, obtendremos: $N_{\text{exp}}^{\text{est}} \geq \lceil \log_2 n \rceil$.

La validez de la desigualdad recíproca se deduce de que la $(0, 1)$ -matriz, cuyas columnas representan la notación de los números naturales $1, 2, \dots, n$ en el sistema binario de numeración, contiene $\lceil \log_2 n \rceil$ filas. El teorema queda demostrado.

Analícemos ahora un caso en que se trata de dos elementos defectuosos. Si para $s = 1$ tenemos la fórmula exacta de $N_{\text{exp}}^{\text{est}}$, entonces cuando $s = 2$, las estimaciones superior e inferior obtenidas no coinciden. Ya en este caso se manifiestan dificultades de los problemas de búsqueda en la planificación del experimento para s grandes.

Al igual que para $s = 1$, al plan estático corresponderá unívocamente una matriz $A = [a_{ij}]$.

En el modelo lineal de la planificación estática obtendremos, como resultado del i -ésimo experimento, una información referente al número de elementos defectuosos que están contenidos en el conjunto Y_i . De este modo, el vector de los resultados será una palabra de longitud N sobre el alfabeto $\{0, 1, 2\}$.

Lema 2. Una matriz $A = [a_{ij}]$ es el plan estático para $s = 2$, es decir, restablece unívocamente el par de elementos defectuosos cuando y sólo cuando son diferentes todas las sumas de columnas cuyas tomadas de dos en dos.

Las sumas de las columnas se entienden como sumas tomadas por componentes, mientras que la sumación será habitual:

$$0 + 0 = 0, 0 + 1 = 1 + 0 = 1, 1 + 1 = 2.$$

La demostración del lema 2 es igual a la del lema 1. La única diferencia consiste en que esta vez el vector de los resultados es una suma de las columnas i -ésima y j -ésima, si (x_i, x_j) es el par de elementos defectuosos.

Teorema 6. En el caso de dos elementos defectuosos

$$N_{\text{exp}}^{\text{est}} \geq \log_3 \binom{n}{2}. \quad (1)$$

Demostración. Supongamos que la $(0, 1)$ -matriz A es de tal género que todas las sumas de las columnas suyas tomadas de dos en dos son diferentes. Por cuanto se pueden comparar en total $\binom{n}{2}$ sumas de las columnas tomadas de dos en dos y cada una de dicha suma es una palabra de longitud N sobre el alfabeto $\{0, 1, 2\}$, donde N es el número de filas en A , entonces resulta válida la desigualdad

$$\binom{n}{2} \leq 3^N$$

Al poner $N = N_{\text{exp}}^{\text{est}}$, obtendremos la afirmación del teorema.

En adelante demostraremos una estimación más fuerte para $N_{\text{exp}}^{\text{est}}$, y por ahora pasaremos a las estimaciones superiores para $N_{\text{exp}}^{\text{est}}$.

Con el fin de obtener estimaciones superiores hagamos uso del método de construcción aleatoria de la matriz del plan estático. Supongamos que la matriz $A = [a_{ij}]$ está construida del modo siguiente:

$$P\{a_{ij} = 1\} = p \geq 0, P\{a_{ij} = 0\} = q \geq 0, p + q = 1;$$

$$(i = 1, \dots, N; j = 1, \dots, n),$$

y que todas las magnitudes aleatorias a_{ij} son independientes en su totalidad.

Veamos un vector columna $\vec{z} = (z_1, \dots, z_n)^T$ de longitud n , compuesto de 2 unidades y $n - 2$ ceros. Este vector se denominará *vector de desarreglo*. En términos de los vectores de desarreglo el lema 2 puede enunciarse del modo siguiente.

Lema 2'. La matriz $A = [a_{ij}]$ es el plan estático, cuando y sólo cuando para cualquier par de diferentes vectores de desarreglo $\vec{z}$ y $\vec{v}$ se verifica la desigualdad: $A\vec{z} \neq A\vec{v}$.

La matriz A se multiplica por el vector de un modo habitual.

Veamos dos vectores de desarreglo $\vec{z}$ y $\vec{v}$ tales que ambas unidades del vector $\vec{z}$ y las unidades del vector $\vec{v}$ ocupan cuatro posiciones diferentes. Calculemos la probabilidad de que $A\vec{z} = A\vec{v}$. Calculemos con anterioridad cuál es la probabilidad de que las i -ésimas componentes de las columnas $A\vec{z}$ y $A\vec{v}$ sean iguales ($i = 1, \dots, N$).

Revisando todos los casos posibles, obtenemos

$$P\{(A\vec{z})_i = (A\vec{v})_i\} = p^4 + 4p^2q^2 + q^4.$$

Por cuanto los elementos de la matriz son magnitudes aleatorias independientes, entonces

$$P\{A\vec{z} = A\vec{v}\} = (p^4 + 4p^2q^2 + q^4)^N.$$

Veamos ahora un par de vectores de desarreglo $\vec{u}$ y $\vec{w}$, en los cuales una componente unitaria coincide, y las segundas son diferentes. Por analogía con el primer caso obtendremos

$$p\{A\vec{u} = A\vec{w}\} = (p^2 + q^2)^N.$$

Estimemos ahora la probabilidad de existencia de los vectores $\vec{z}_1$ y $\vec{z}_2$ tales que $A\vec{z}_1 = A\vec{z}_2$:

$$P\{\exists \vec{z}_1, \vec{z}_2: A\vec{z}_1 = A\vec{z}_2\} \leq M_1(p^4 + 4p^2q^2 + q^4)^N + M_2(p^2 + q^2)^N,$$

donde M_1 es el número de pares de vectores de primer tipo, y M_2 , el número de pares de vectores de segundo tipo.

Los razonamientos combinatorios sencillos conducen a las siguientes igualdades:

$$M_1 = 3 \binom{n}{4}, \quad M_2 = 3 \binom{n}{3}.$$

De este modo,

$$P\{\exists \vec{z}_1, \vec{z}_2: A\vec{z}_1 = A\vec{z}_2\} \leq 3 \binom{n}{4} (p^4 + 4p^2q^2 + q^4)^N + 3 \binom{n}{3} (p^2 + q^2)^N.$$

Supongamos que para ciertos p y q ($p, q \geq 0$, $p + q = 1$) el número $\tilde{N} = \tilde{N}(n)$ es tal que

$$3 \binom{n}{4} (p^4 + 4p^2q^2 + q^4)^{\tilde{N}} + 3 \binom{n}{3} (p^2 + q^2)^{\tilde{N}} < 1. \quad (2)$$

De esto se deducirá que con la probabilidad superior a cero existe una matriz A de dimensión $\tilde{N} \times n$, para la cual la desigualdad

$$A\vec{z} \neq A\vec{v}$$

se verifica para cualesquiera dos diferentes vectores de desarreglo $\vec{z}$ y $\vec{v}$. Así pues, la matriz A corresponderá al plan estático.

Por consiguiente, $N_{\text{est}}^{\text{est}} \leq \tilde{N}$.

Como probabilidades p y q podemos elegir cualesquiera números no negativos tales que $p + q = 1$. Sin embargo, resulta más ventajoso tomar tales p y q con los que las expresiones $p^4 + 4p^2q^2 + q^4$ y $p^2 + q^2$ sean las más pequeñas posibles. Las búsquedas no complejas de los valores extremos conducen a los siguientes resultados:

$$\min_{\substack{p, q \geq 0 \\ p + q = 1}} (p^4 + 4p^2q^2 + q^4) = \frac{3}{8} \quad \text{y se consigue para } p = q = \frac{1}{2}.$$

$\min_{\substack{p, q \geq 0 \\ p+q=1}} (p^2 + q^2) = \frac{1}{2}$ y se consigue para $p = q = \frac{1}{2}$.

Pongamos, por eso, $p = q = \frac{1}{2}$ y reescribamos la relación (2):

$$3 \binom{n}{4} \binom{3}{8}^N + 3 \binom{n}{3} \binom{1}{2}^N < 1. \quad (3)$$

Si se verifican simultáneamente dos desigualdades:

$$\begin{aligned} 3 \binom{n}{4} \binom{3}{8}^N &< \frac{1}{2}, \\ 3 \binom{n}{3} \binom{1}{2}^N &< \frac{1}{2}, \end{aligned} \quad (4)$$

la desigualdad (3) también se cumplirá. No es difícil comprobar que las desigualdades (4) son válidas para $N = \lceil 3 \log_2 n \rceil$. De este modo, hemos demostrado el teorema siguiente.

Teorema 7. En el caso de dos elementos defectuosos

$$N_{\text{exp}}^{\text{est}} \leq \lceil 3 \log_2 n \rceil.$$

Observación. El resultado aducido se ha obtenido de una manera que no puede considerarse constructiva, pues no hemos construido una matriz concreta del plan estático de dimensión $\lceil 3 \log_2 n \rceil \times n$, sino que sólo demostramos la existencia de la misma.

Ha de notarse, sin embargo, que en dicho problema la victoria queda por el lado de los "constructivistas". El plan estático con el número de filas $2 \cdot \lceil \log_2 n \rceil$ (véase [83]), construido con ayuda de los códigos de Bose nos da la siguiente estimación:

$$N_{\text{exp}}^{\text{est}} \leq 2 \lceil \log_2 n \rceil.$$

No obstante, para s grandes de construcción aleatoria de una matriz del plan estadístico proporciona resultados considerablemente mejores que el método de los "constructivistas" (véase [84]).

Empleando métodos análogos podemos obtener también estimaciones superiores de la longitud del plan estadístico para el llamado modelo disyuntivo, en el cual tenemos, como resultado de la i -ésima comprobación, sólo información referente a la presencia de por lo menos un elemento defectuoso en el conjunto Y_i .

7.3. MÉTODO DE ENTROPÍA

Este método permite mejorar la estimación (1) obtenida en el teorema 6. Recordemos previamente la definición y propiedades principales de la entropía.

Sea una distribución finita de las probabilidades $\mathcal{P} = \{p_1, \dots, p_m\}$,

$\sum_{i=1}^m p_i = 1$, $p_i \geq 0$, $i = 1, \dots, m$. Se denominará entropía de la distribución $\mathcal{P}$ a la magnitud $H(\mathcal{P}) = - \sum_{i=1}^m p_i \log_2 p_i$. Además, si alguna p_i es nula, entonces, en lugar de un sumando indeterminado $-p_i \log_2 p_i$ se pone en la suma el cero (lo que es natural, puesto que $\lim_{x \rightarrow 0} x \log_2 x = 0$).

Examinemos las propiedades principales del concepto introducido.

Propiedad 1. Se verifica la desigualdad

$$H(\mathcal{P}) \geq 0,$$

donde la igualdad se consigue cuando y sólo cuando una de las probabilidades p_i es igual a la unidad, y las demás, a cero.

Propiedad 2. Se verifica la desigualdad:

$$H(\mathcal{P}) \leq \log_2 m, \quad (5)$$

en la cual la igualdad se consigue cuando y sólo cuando $p_1 = \dots = p_m = 1/m$.

En efecto,

$$H(\mathcal{P}) - \log_2 m = \sum_{i=1}^m p_i \log_2 (1/p_i m) \leq \log_2 e \sum_{i=1}^m p_i \left(\frac{1}{p_i m} - 1 \right) = 0.$$

Aquí se ha aprovechado la desigualdad $\log_2 x \leq (x-1)\log_2 e$, en la cual el signo de igualdad se consigue sólo para $x = 1$. Por consiguiente, la igualdad en (5) tiene lugar cuando y sólo cuando $p_1 = \dots = p_m = 1/m$.

Veamos la cuestión sobre la entropía de la distribución compleja de las probabilidades $\mathcal{P} = \{p_{ij}\}$, $i = 1, \dots, m_1$; $j = 1, \dots, m_2$; $p_{ij} \geq 0$, $\sum_{i,j} p_{ij} = 1$. Formemos dos distribuciones simples de las probabilidades $\mathcal{A}_1$ y $\mathcal{A}_2$ del modo siguiente:

$$\begin{aligned} \mathcal{A}_1 &= \{p_i^{(1)}\}, p_i^{(1)} = \sum_{j=1}^{m_2} p_{ij}, i = 1, \dots, m_1; \\ \mathcal{A}_2 &= \{p_j^{(2)}\}, p_j^{(2)} = \sum_{i=1}^{m_1} p_{ij}, j = 1, \dots, m_2; \end{aligned}$$

Está claro que $\mathcal{A}_1$ y $\mathcal{A}_2$ son realmente distribuciones de las probabilidades, es decir, $p_i^{(1)} \geq 0$, $p_j^{(2)} \geq 0$, $i = 1, \dots, m_1$; $j = 1, \dots, m_2$; $\sum_{i=1}^{m_1} p_i^{(1)} = 1$, $\sum_{j=1}^{m_2} p_j^{(2)} = 1$. La propiedad siguiente se refiere a la relación existente entre

las magnitudes $H(\mathcal{P})$, $H(\mathcal{A}_1)$ y $H(\mathcal{A}_2)$.

Propiedad 3 (subaditividad de la entropía). Se verifica la desigualdad:

$$H(\mathcal{P}) \leq H(\mathcal{A}_1) + H(\mathcal{A}_2),$$

donde el signo de igualdad tiene lugar cuando y sólo cuando las distribuciones $\mathcal{A}_1$ y $\mathcal{A}_2$ son independientes, es decir, para cualesquiera i, j se cumple la relación: $p_{ij} = p_i \cdot p_j$.

En efecto, $H(\mathcal{A}) - H(\mathcal{A}_1) - H(\mathcal{A}_2) = \sum_{i,j} p_{ij} \log_2 \frac{p_i^{(1)} p_j^{(2)}}{p_{ij}}$. Aplican-

do la desigualdad $\log_2 x \leq (x - 1) \log_2 e$, obtenemos para cada término de la suma en el segundo miembro la siguiente estimación:

$$p_{ij} \log_2 \frac{p_i^{(1)} p_j^{(2)}}{p_{ij}} \leq p_{ij} \left(\frac{p_i^{(1)} p_j^{(2)}}{p_{ij}} - 1 \right) \log_2 e,$$

con la particularidad de que la igualdad tiene lugar cuando y sólo cuando $p_{ij} = p_i^{(1)} \cdot p_j^{(2)}$. Por consiguiente,

$$H(\mathcal{A}) - H(\mathcal{A}_1) - H(\mathcal{A}_2) \leq \log_2 e \left(- \sum_{i,j} p_{ij} + \sum_{i,j} p_i^{(1)} p_j^{(2)} \right) = 0,$$

lo que se trataba de demostrar.

Hemos de notar que la propiedad de subaditividad de la entropía puede extenderse con facilidad a un número arbitrario de distribuciones $\mathcal{A}_1, \mathcal{A}_2, \dots, \mathcal{A}_n$.

Ejemplo. Sea $1 \leq k \leq n/2$, $h(p) = -p \log_2 p - (1-p) \log_2 (1-p)$.

Demostremos que $\sum_{i=0}^k \binom{n}{i} \leq 2^{nh(k/n)}$.

Hagamos $p = k/n$. Tenemos $\binom{n-1}{i} \binom{n}{i+1}^{-1} = \frac{i+1}{n} \leq p$

para $i = 0, \dots, k-1$. Por consiguiente, $\tilde{p} = \left(\sum_{i=0}^{k-1} \binom{n-1}{i} \right) \left(\sum_{i=0}^k \binom{n}{i} \right)^{-1} < p$. Sea $i_j \in [0, 1]$, $j = 1, \dots, n$,

$$p_{i_1, \dots, i_n} = \begin{cases} 0, & \text{si } \sum_{j=1}^n i_j > k, \\ \left(\sum_{i=0}^n \binom{n}{i} \right)^{-1}, & \text{si } \sum_{j=1}^n i_j \leq k, \end{cases}$$

$$P^*, \dots, *, i_j^*, \dots, * = \begin{cases} \tilde{p}, & \text{si } i_j = 1 \\ 1 - \tilde{p}, & \text{si } i_j = 0, \end{cases} \text{ así que}$$

$$p^*, \dots, *, i_j^*, \dots, * = \sum_{i_1, \dots, i_{j-1}, i_{j+1}, \dots, i_n} p_{i_1, \dots, i_n}.$$

Aprovechando la propiedad de subaditividad de la entropía, obtenemos

$$\begin{aligned}\log_2 \sum_{i=0}^k \binom{n}{i} &= - \sum_{i_1, \dots, i_n} p_{i_1, \dots, i_n} \log_2 p_{i_1, \dots, i_n} = \\ &= H(\{p_{i_1, \dots, i_n}\}) \leq \sum_{j=1}^n H(\{p_{*, \dots, *, i_j, *, \dots, *}\}) = \\ &= nh(\bar{p}) < nh(p) \quad \left(\text{puesto que } \bar{p} < p \leq \frac{1}{2} \right),\end{aligned}$$

lo que se requería demostrar.

Volvamos al problema de experimentos de rechazo. Examinemos el caso de un modelo lineal del plan estático en que se tienen dos elementos defectuosos.

Teorema 8.

$$N_{\text{exp}}^{\text{est}} \geq (1 + o(1)) \cdot \frac{3}{4} \log_2 n.$$

Demostración. Sea $\vec{z}$ un vector columna de desarreglo que contiene dos unidades y $(n-2)$ ceros, y sea $\vec{e}$ una columna de resultados de los experimentos, mientras que A representa una matriz del plan estático. Entonces

$$\vec{e} = A\vec{z}.$$

Introduzcamos en el conjunto Ω de vectores de desarreglo una distribución uniforme $\mathcal{P}$: $P(\vec{z}) = \binom{n}{2}^{-1}$. Siendo la aplicación $A: \Omega \rightarrow A(\Omega)$ biunívoca, en el conjunto $A(\Omega)$ se induce la distribución uniforme de probabilidades Q :

$$P(\vec{e}) = \binom{n}{2}^{-1} \quad \text{y} \quad H(Q) = H(\mathcal{P}) = \log_2 \binom{n}{2}.$$

Supongamos que en la i -ésima fila de la matriz A se tienen $n_0^{(i)}$ ceros y $n_1^{(i)}$ unidades. La distribución $\mathcal{P}$ induce en el conjunto $\{0, 1, 2\}^N$ distribuciones Q_i que se determinan del modo siguiente:

$$P_0^{(i)} = P\{e_i = 0\} = \binom{n_0^{(i)}}{2} \binom{n}{2}^{-1} = \frac{n_0^{(i)} \cdot (n_0^{(i)} - 1)}{n \cdot (n-1)},$$

$$P_1^{(i)} = P\{e_i = 1\} = \binom{n_0^{(i)}}{1} \binom{n_1^{(i)}}{1} \cdot \binom{n}{2}^{-1} = \frac{2n_0^{(i)} \cdot n_1^{(i)}}{n \cdot (n-1)},$$

$$P_2^{(i)} = P\{e_i = 2\} = \binom{n_1^{(i)}}{2} \cdot \binom{n}{2}^{-1} = \frac{n_1^{(i)} \cdot (n_1^{(i)} - 1)}{n \cdot (n-1)}.$$

Está claro que $P_i^{(i)} = \sum_{\vec{z} \in A(\Omega): e_i = i} P(\vec{z})$. Recurriendo a la propiedad de subaditividad de la entropía, obtenemos: $\log_2 \binom{n}{2} = H(Q) \leq \sum_{i=1}^N H(Q_i)$. Nos

convencemos inmediatamente de que $H(Q_i) = \varphi(x_i) \cdot (1 + o(1))$, donde $x_i = \frac{n_0^{(i)}}{n}$, $n \rightarrow \infty$, y $\varphi(x) = -x^2 \log_2 x^2 - 2x(1-x) \log_2 x(1-x) - (1-x)^2 \log_2 (1-x)^2$.

Por cuanto

$$\max_{x \in [0,1]} \varphi(x) = \varphi\left(\frac{1}{2}\right) = \frac{3}{2} \quad \text{y} \quad \sum_{i=1}^N H(Q_i) \leq (1 + o(1)) N \max_{x \in [0,1]} \varphi(x),$$

entonces

$$N \geq \frac{\log_2 \binom{n}{2}}{\max_{x \in [0,1]} \varphi(x)} \cdot (1 + o(1)) = (1 + o(1)) \cdot \frac{4}{3} \log_2 n.$$

El teorema está demostrado.

Haciendo uso de las ideas análogas (pero en otro lenguaje matemático) y aplicando adicionalmente razonamientos combinatorios más finos, Lindström [85] obtuvo una estimación más fuerte:

$$N_{\text{exp}}^{\text{est}} \geq (1 + o(1)) \cdot \frac{5}{3} \log_2 n.$$

7.4. MÉTODO DE BALANCE ALEATORIO

En los párrafos antecedentes se analizaba el modelo matemático del problema de experimentos de rechazo, en el que se necesitaba hallar s elementos fijos, pero desconocidos de antemano, en un conjunto X compuesto de n elementos. Ahora suponemos que a todo elemento $x \in X$ se le ha hecho corresponder cierto número real $\omega(x)$, llamado *peso* del elemento citado, y entre los elementos del conjunto X sólo algunos, que no se conocen de antemano y cuya cantidad es igual a $s \geq 2$, poseen peso no nulo. Se requiere separar estos s elementos del conjunto X y determinar los pesos de ellos. Se permite en cada experimento elegir cierto subconjunto $Y \subseteq X$ y determinar la suma de pesos de los elementos pertenecientes a Y (es decir, "ponderar" los elementos de Y). Hace falta planificar los experimentos de una manera tal que a base de sus resultados se puedan determinar los pesos $\omega_j = \omega(x_j)$ de los elementos, pertenecientes a X ($j = 1, 2, \dots, n$).

En lo que sigue nos adheriremos a la siguiente terminología: los pesos $\omega_1, \dots, \omega_n$ de los elementos $x_1, \dots, x_n$ se llamarán *factores*; el factor ω_j se denominará *significativo*, si $\omega_j \neq 0$. El número s de factores significativos se supone pequeño en comparación con el número total de factores n .

Partiendo de la reunión de los subconjuntos $Y_i \subseteq X$ ($i = 1, \dots, N$) construyamos una $(N \times n)$ -matriz de los experimentos $A = [a_{ij}]$, donde $a_{ij} = 1$, siempre que $x_j \in Y_i$, y $a_{ij} = 0$, en el caso contrario. Así pues, en cada i -ésimo experimento ($i = 1, \dots, N$) se halla la magnitud $e_i = \sum_{j=1}^n a_{ij} \omega_j$.

La matriz de los experimentos $A = [a_{ij}]$ se denomina, además, plan de los experimentos. El plan ha de satisfacer los siguientes requisitos principales: sencillez en la realización y economía. El cumplimiento de las condiciones mencionadas conduce a la necesidad de minimizar el número N de experimentos en el plan.

En el párrafo anterior se ha considerado el caso cuando $s = 2$ y todos los factores significativos son iguales a uno. En cambio, si el número s de factores significativos es arbitrario, no será difícil mostrar que para su separación unívoca del número total de factores n harán falta no menos de

$\log_{s+1} \binom{n}{s} \sim \frac{s}{\log_2 s} \log_2 n$ experimentos ($s \leq n$). L. D. Meshalkin fue el primero en acentuar que los factores significativos son diferentes en la práctica e, incluso, poseen cierto carácter de inconmensurabilidad (véase [86]). Esta circunstancia hace posible construir los planes que permiten determinar con eficacia dichos factores en el transcurso de $\log_2 n$ experimentos, lo que es considerablemente inferior a la estimación inferior aducida.

Definición. Los factores $\omega_1, \dots, \omega_n$ se llaman *inconmensurables* sobre el conjunto $\Lambda \subseteq \mathbb{Z}$ (donde $0 \in \Lambda$), si de la condición $\sum_{j=1}^n \lambda_j \omega_j = 0$, donde $\lambda_j \in \Lambda$, se deduce que $\lambda_j \omega_j = 0$ para cualesquiera $j = 1, \dots, n$, es decir, $\lambda_j = 0$ para todos los factores significativos ω_j .

El problema planteado se resolverá para $\Lambda = \mathbb{Z}$, por el método de balance aleatorio. He aquí el esquema que debe seguirse. Se elabora un algoritmo, para el cual de los datos de partida sirve una $\{0, 1\}$ -matriz $A = [a_{ij}]$ que contiene N filas y n columnas, y, además, un vector columna $\vec{e} = (e_1, \dots, e_N)^T \in \mathbb{R}^N$ (T es la operación de transposición). Si la matriz y el vector $\vec{e}$ poseen ciertas propiedades, las cuales trataremos más abajo, el algoritmo elabora el vector $\vec{\omega}' \in \mathbb{R}^n$; de lo contrario el algoritmo no nos da nada. Este algoritmo resulta ser de tal índole que de $\vec{e} = A\vec{\omega}$ se deduce: $\vec{\omega}' = \vec{\omega}$. Se demuestra que para todo $\beta \in]0, 1[$ se puede indicar tal número $N = N(n, s, \beta)$ que la matriz A , construida al azar, y el vector $\vec{e} = A\vec{\omega}$ posean las propiedades indispensables con la probabilidad no inferior a $1 - \beta$.

Ha de indicarse que el número s de factores significativos no es obligatoriamente conocido de antemano y puede determinarse en el transcurso de los experimentos.

Teorema 9. Supongamos que los factores $\omega_1, \dots, \omega_n$ son inconmensurables sobre el conjunto $\mathbb{Z}$, y el plan de los experimentos $A = [a_{ij}]$ se construye de un modo aleatorio:

$$P[a_{ij} = 0] = P[a_{ij} = 1] = \frac{1}{2}; \quad i = 1, \dots, N; \quad j = 1, \dots, n.$$

Entonces, si s es el número desconocido de factores significativos, y β , un número real, con la particularidad de que $0 < \beta < 1$, entonces para

$$N > s + \log_2(n - s + 1) - \log_2 \beta$$

podemos determinar, con la probabilidad no inferior a $1 - \beta$, todos los factores significativos y su número s . En este caso el error en la determinación de los factores está excluido.

Observemos que si el número s de factores significativos no sobrepasa cierto número conocido de antemano k : $s \leq k \ll n$, el teorema citado fija la frontera para el número de experimentos que han de ser realizados, para que el plan aleatorio construido determine los factores significativos con una probabilidad no inferior a la prefijada.

Previamente a la demostración del teorema aduzcamos el siguiente lema de L. D. Meshalkin.

Lema 3. Cualquier plano l -dimensional en un espacio real n -dimensional ($2 \leq n$; $l \leq n$) contiene no más de 2^l vértices de un cubo n -dimensional.

La demostración se realizará por inducción respecto de n . Cuando $n = 2$ l puede igualarse a 1 ó a 2, y la afirmación del lema resulta evidente. Supongamos que dicha afirmación ya está demostrada para todos los números $n' < n$, donde $n \geq 3$. Demostremosla también para $n' = n$.

Elijamos un sistema de coordenadas en el espacio R^n de un modo tal que el cubo n -dimensional dado se haga unitario, es decir, que las coordenadas de sus vértices sean iguales a 0 ó 1. Está claro que si $l = n$, cada plano l -dimensional puede contener no más de 2^l vértices del cubo (número total de todos los vértices suyos). Si $l = 1$, entonces la recta $\bar{x}(t) = \bar{a} + (\bar{b} - \bar{a})t$, que pasa por diferentes vértices $\bar{a} = (a_1, \dots, a_n)^T$ y $\bar{b} = (b_1, \dots, b_n)^T$ ($a_i, b_i \in [0, 1]$, $i = 1, \dots, n$), no puede contener ningún otro vértice más. En efecto, sea i tal que $a_i \neq b_i$. Entonces, $x_i(0) = x_i(1)$, y $x_i(0), x_i(1) \in \{0, 1\}$. Por consiguiente, para todos los $t \neq 0, 1$:

$$x_i(t) = a_i + b_i - a_i)t \notin \{0, 1\}.$$

Sea ahora l tal que $1 < l < n$. Denotemos $R_i^n = \{\bar{x} \in R^n : x_n = i\}$, $i = 0, 1$. Si un plano l -dimensional L es paralelo a R_0^n o a R_1^n y no yace en estos últimos, entonces el plano no contiene ningún vértice del cubo. En cambio, si L está situado en uno de los hiperplanos R_0^n ó R_1^n , la afirmación del lema se cumplirá por la hipótesis de la inducción. Resta analizar un caso en que L interseca ambos hiperplanos. Designemos $L_i = L \cap R_i^n$, $i = 0, 1$. Está claro que las dimensiones de los planos L_0 y L_1 son iguales a $l - 1$. Los vértices del cubo, situados en cada uno de los hiperplanos R_i^n ($i = 0, 1$), forman un cubo $(n - 1)$ -dimensional y, por consiguiente, por hipótesis de la inducción, el plano L_i contiene 2^{l-1} tales vértices a lo sumo. Por eso, el plano L contiene no más de $2^{l-1} + 2^{l-1} = 2^l$ vértices del cubo n -dimensional dado. El lema queda demostrado.

Demostración del teorema 9. Introduzcamos las siguientes designaciones: $\bar{a}_j$ es la j -ésima columna de la matriz $A = [a_{ij}]$; $L(\bar{a}_1, \dots, \bar{a}_m)$ es la cápsula lineal de los vectores $\bar{a}_1, \dots, \bar{a}_m$. Diremos que la matriz A y el vector $\bar{e}$ poseen la propiedad S , si en A existen columnas linealmente

independientes $\vec{a}_{j_1}, \dots, \vec{a}_{j_m}$, cuya cápsula lineal no contiene ninguna otra columna y el sistema de ecuaciones

$$\vec{e} = \sum_{r=1}^m \vec{a}_{j_r} x_r \quad (6)$$

tiene una única solución. Si la matriz A y el vector $\vec{e}$ poseen la propiedad S , entonces, al resolver el sistema (6), formamos un vector $\vec{\omega}' = (\omega'_1, \dots, \omega'_n)^T$, haciendo

$$\omega'_j = \begin{cases} \bar{x}_r & \text{para } j = j_r \ (r = 1, \dots, m), \\ 0 & \text{para } j \notin \{j_1, \dots, j_m\}, \end{cases}$$

donde $(\bar{x}_1, \dots, \bar{x}_m)$ es la solución del sistema (6). Observemos que si el juego de números de las columnas $j_1, \dots, j_m$ en la propiedad S se define de una manera no unívoca, entonces puede elegirse arbitrariamente. Por consiguiente, el vector construido $\vec{\omega}'$ es precisamente vector buscado $\vec{\omega}$, si $\vec{e} = A\vec{\omega}$. Para demostrar esto, basta comprobar que los números de todos los factores significativos ω_j pertenecen al conjunto $\{j_1, \dots, j_m\}$. En efecto, si esto no es así, se encontraría un juego $\{\alpha_1, \dots, \alpha_l\}$ de números de las columnas que no está íntegramente contenido en el conjunto citado y que satisface el sistema de ecuaciones:

$$\vec{e} = \sum_{r=1}^l a_{\alpha_r} \omega_{\alpha_r},$$

donde $\omega_{\alpha_1}, \dots, \omega_{\alpha_l}$ son todos factores significativos.

Veamos un sistema de ecuaciones

$$\sum_{r=1}^m \vec{a}_{j_r} x_r - \sum_{r=1}^l \vec{a}_{\alpha_r} y_r = 0 \quad (7)$$

con las incógnitas $x_1, \dots, x_m, y_1, \dots, y_l$. De la definición de la propiedad S y de la elección del juego $\{\alpha_1, \dots, \alpha_l\} \not\subseteq \{j_1, \dots, j_m\}$ se deduce que el rango del sistema no es inferior a $m+1$. Por consiguiente, haciendo uso del procedimiento de Gauss, podemos excluir de dicho sistema de ecuaciones las variables $x_1, \dots, x_m$ y obtener por lo menos una ecuación que contenga sólo las variables y_r ($r = 1, \dots, l$):

$$0 = \sum_{r=1}^l b_r y_r, \quad \sum_{r=1}^l b_r^2 \neq 0. \quad (8)$$

Los coeficientes de las incógnitas en el sistema (7) eran enteros, razón por la cual pueden considerarse enteros también los coeficientes b_r en (8). El sistema de ecuaciones (7) tiene la solución: $x_r = \bar{x}_r$, $r = 1, \dots, m$; $y_r = \omega_{\alpha_r}$, $r = 1, \dots, l$. Por eso, el juego $\omega_{\alpha_1}, \dots, \omega_{\alpha_l}$ forma la solución de la ecuación (8), lo que contradice la condición de inconmensurabilidad de los factores significativos $\omega_{\alpha_1}, \dots, \omega_{\alpha_l}$.

Se ha mostrado, pues, que si la matriz A y el vector $\vec{e}$ poseen la propiedad S , entonces el vector desconocido $\vec{\omega}$ queda unívocamente definido. Estimemos la probabilidad P de que construyendo al azar un plan con el

número de experimentos igual a $N > s + \log_2(n - s + 1) - \log_2 \beta$, el par $(A, \bar{e})$ poseerá la propiedad S , donde $\bar{e} = A\bar{\omega}$.

Designemos mediante $j_1, \dots, j_s$ los números desconocidos de los factores significativos. Estudiemos los sucesos:

$$\begin{aligned} X &= \{\dim L(\bar{a}_{j_1}, \dots, \bar{a}_{j_s}) = s\}, \\ Y_j &= \{\bar{a}_j \notin L(\bar{a}_{j_1}, \dots, \bar{a}_{j_s})\}, j \notin \{j_1, \dots, j_s\}, \\ Y &= \bigcap_{j \notin \{j_1, \dots, j_s\}} Y_j, Z = X \cap Y. \end{aligned}$$

Está claro que la aparición del suceso Z lleva consigo el que el par $(A, \bar{e})$, donde $\bar{e} = A\bar{\omega}$, posee la propiedad S . Por consiguiente, $P \geq P(Z)$. Estimemos la probabilidad del suceso Z . Tenemos

$$P(Z) = P(X \cap Y) = P(X) \cdot P(Y|X).$$

Ciertos razonamientos combinatorios no complejos con la aplicación del lema 3 nos conducen a una desigualdad

$$P(Y|X) \geq (1 - 2^{-N+s})^{n-s}.$$

Para poder estimar la probabilidad del suceso X analicemos los siguientes sucesos:

$$\begin{aligned} X_1 &= \{\dim L(\bar{a}_{j_1}) = 1\}, X_2 = \{\dim L(\bar{a}_{j_1}, \bar{a}_{j_2}) = 2\}, \dots, \\ X_{s-1} &= \{\dim L(\bar{a}_{j_1}, \dots, \bar{a}_{j_{s-1}}) = s-1\}, \\ \bar{X}_1 &= \{\bar{a}_{j_1} \notin L(\bar{a}_{j_1})\}, \bar{X}_2 = \{\bar{a}_{j_2} \notin L(\bar{a}_{j_1}, \bar{a}_{j_2})\}, \dots, \\ \bar{X}_{s-1} &= \{\bar{a}_{j_{s-1}} \notin L(\bar{a}_{j_1}, \dots, \bar{a}_{j_{s-1}})\}. \end{aligned}$$

Al hacer uso del lema 3, no es difícil notar que

$$P(X_1) = 1 - 2^{-N}, P(\bar{X}_r|X_r) \geq 1 - 2^{-N+r}, r = 1, \dots, s-1.$$

Por eso, $P(X) = P(X_1) \prod_{r=1}^{s-1} P(\bar{X}_r|X_r) \geq \prod_{r=0}^{s-1} (1 - 2^{-N+r}) \geq 1 - 2^{-N+s}$.

para $s \geq 2$. La demostración de la última desigualdad en esta cadena puede fácilmente obtenerse por inducción respecto de s . Así pues, $P(Z) \geq (1 - 2^{-N+s})^{n-s+1} > 1 - (n-s+1)2^{-N+s}$ para $N > s$, puesto que para $0 < \alpha < 1$ es válida la desigualdad: $(1 - \alpha)^m > 1 - m\alpha$. Por consiguiente, para $N > s + \log_2(n - s + 1) - \log_2 \beta$ se verifica la desigualdad: $P(Z) > 1 - (n - s + 1)2^{-N+s} > 1 - \beta$, lo que finaliza la demostración del teorema.

M. B. Maljútov y M. S. Pínsker [87] han estudiado el problema aducido para las condiciones más generales, cuando los factores $\omega_1, \dots, \omega_n$ se suponen incommensurables sobre el conjunto $\Lambda = [-1, 0, 1]$ (los experimentos se realizan en este caso no de un modo aleatorio, sino sucesivamente, o sea, tomando en consideración los resultados anteriores. Con ellos se ha demostrado el teorema siguiente.

Teorema 10. Supongamos que los factores $\omega_1, \dots, \omega_n$ son inconmen-

surables sobre el conjunto $\Lambda = \{-1, 0, 1\}$; y sea s un número desconocido de factores significativos. Existe, entonces, un método no aleatorio de realizar los experimentos que permite determinar ciertamente todos los factores en el transcurso de no más de $s \log_2 s + \log_2 n$ experimentos. El número s se halla en este caso en el proceso de los experimentos.

Demostración. Con ayuda del primer experimento determinamos $e_1 = \sum_{j=1}^n \omega_j$. Los experimentos ulteriores se realizan de un modo inductivo.

Supongamos que tras cierta cantidad de experimentos el conjunto de todos los factores haya resultado partido en $r_m + 1$ subconjuntos G_i^m , $i = 0, 1, \dots, r_m$ ($1 \leq r_m \leq s$), con la particularidad de que G_0^m no contiene factores significativos, mientras que todos los G_i^m restantes contienen cada uno por lo menos un factor significativo, y se conocen las sumas $\sigma_i^m = \sum_{\omega \in G_i^m} \omega_j$,

$i = 1, \dots, r_m$. En el experimento siguiente partiremos cada uno de los subconjuntos G_i^m ($i = 1, \dots, r_m$) en partes aproximadamente iguales: $G_i^m = G_{i0}^m \cup G_{i1}^m$, $G_{i0}^m \cap G_{i1}^m = \emptyset$, $|G_{i0}^m| - |G_{i1}^m| \leq 1$, y hallaremos la suma $c = \sum_{i=1}^{r_m} \sum_{\omega \in G_{i0}^m} \omega_j$. Si $e \in \left\{ \sum_{i=1}^{r_m} \varepsilon_i \sigma_i^m \mid \varepsilon_i \in \{0, 1\} \right\}$, entonces, haciendo uso de

la condición de inconmensurabilidad de los factores, llegamos a la conclusión de que cada uno de los subconjuntos $G_{i\varepsilon_i}^m$ contiene sólo factores insignificativos. Entonces, al elegir $G_i^{m+1} = G_{i\varepsilon_i}^m$, $\sigma_i^{m+1} = \sigma_i^m$ ($i = 1, \dots, r_m$), $r_{m+1} = r_m$, $G_0^{m+1} = \left(\bigcup_{i=1}^{r_m} G_{i\varepsilon_i}^m \right) \cup G_0^m$, donde $\varepsilon_i = (c_i + 1) \bmod 2$,

continuaremos los experimentos según el esquema citado hasta que queden determinados todos los factores significativos, es decir, hasta que r_m no se haga igual a s , y cada uno de los subconjuntos G_i^m no se componga de un único factor significativo. En cambio, si $e \in \left\{ \sum_{i=1}^{r_m} \varepsilon_i \sigma_i^m \mid \varepsilon_i \in \{0, 1\} \right\}$,

tendremos que realizar un ciclo de experimentos para determinar las sumas $\sigma_{i\varepsilon}^m = \sum_{\omega \in G_{i\varepsilon}^m} \omega_j$ de los factores significativos en todas las mitades $G_{i\varepsilon}^m$,

$i = 1, \dots, r_m$; $\varepsilon = 0, 1$. Con este fin realicemos un experimento y determinemos la suma $\sum_{i=1}^{\lfloor r_m/2 \rfloor} \sigma_{i0}^m$. Aclaremos si dicha suma pertenece al conjunto $\left\{ \sum_{i=1}^{\lfloor r_m/2 \rfloor} \varepsilon_i \sigma_i^m \mid \varepsilon_i \in \{0, 1\} \right\}$. Si la respuesta resulta positiva, llegamos a una

conclusión de que cada uno de los subconjuntos $G_{i\varepsilon_i}$, $i = 1, \dots, \left\lfloor \frac{r_m}{2} \right\rfloor$ contiene sólo factores insignificativos. De lo contrario, realizamos un experimento más y determinamos la suma $\sum_{i=1}^{\lfloor r_m/4 \rfloor} \sigma_{i0}^m$, etc. Indiquemos que en este

caso las sumas complementarias, tales como $\sum_{i \in [r_m/2]+1}^{r_m} \sigma_{i0}^m$, se determinan automáticamente y en adelante han de tratarse análogamente, dividiendo, si es necesario, el conjunto en el que viene definido el índice de sumación, aproximadamente por la mitad. Así pues, como resultado de tal ciclo de experimentos, determinamos todas las sumas $\sigma_{ic}^m = \sum_{j \in G_{ic}^m} w_j$, $i = 1, \dots, r_m$;

$\varepsilon = 0, 1$, y aclaramos cuáles de los subconjuntos G_{ic}^m se componen íntegramente sólo de los factores significativos. Al relacionar todos estos subconjuntos, junto con los subconjuntos G_0^m , al subconjunto G_0^{m+1} y al reenumerar todos los subconjuntos nuevamente mediante los símbolos $G_1^{m+1}, \dots, G_{r_{m+1}}^{m+1}$, donde $r_{m+1} > r_m$, llegamos a la misma situación que teníamos al principio de nuestros razonamientos.

Estimemos el número de experimentos del plan construido sucesivamente. No es difícil notar que en cada ciclo el número de experimentos no sobrepasa de $(r_{m+1} - r_m)(1 + \log_2 r_m)$. Cada uno de los demás experimentos que no integran los ciclos disminuye el número de factores, que se supone son significativos, en dos veces (con una corrección evidente por el hecho de que este número puede no dividirse por 2). Por consiguiente, el número total N de experimentos del plan no sobrepasa $s \log_2 s + \log_2 n$:

$$N \leq 1 + \log_2 n + \sum_m (r_{m+1} - r_m)(1 + \log_2 r_m) \leq 1 + \\ + \log_2 n + \sum_m \sum_{i: r_m \leq i < r_{m+1}} (1 + \log_2 i) = 1 + \log_2 n + \sum_{i=1}^{s-1} (1 + \log_2 i) \quad (\text{puesto que } 1 = r_1 < r_2 < \dots < r_m < r_{m+1} < \dots < r_p = s).$$

Por consiguiente,

$$N \leq s + \log_2 n + \sum_{i=2}^{s+1} \log_2 i \leq s + \log_2 n + \int_2^s \log_2 x dx \leq s \log_2 s + \log_2 n.$$

El teorema queda demostrado.

7.5. SISTEMAS SEPARADORES DE SUBCONJUNTOS

Recordemos que el sistema de subconjuntos $\{Y_1, Y_2, \dots, Y_N\}$ de un conjunto finito X lleva el nombre de sistema *separador*, si en él para cualesquiera dos elementos distintos de X existe Y_i ($1 \leq i \leq N$) que contiene un solo elemento de los dos mencionados (véase § 1.5). Pongamos $n = |X|$. En este párrafo se analiza el problema de buscar un sistema de separación mínimo, es decir, un sistema compuesto de un número mínimo posible $N = N(n, k)$ de subconjuntos, a condición de que cada subconjunto contiene exactamente k elementos. Nos será de interés también un sistema separador mínimo, cada subconjunto del cual se compone de no más k

elementos. La cantidad de subconjuntos de tal sistema la designamos con $N(n, \leq k)$.

Al concepto de sistema separador llevan algunos problemas de planificación de los experimentos de rechazo (véase § 7.2). Supongamos que en nuestra disposición se tiene una balanza analítica, con ayuda de la cual se pueden determinar los pesos de los subconjuntos del conjunto X , que contiene n monedas. Se conoce que en el proceso de un solo experimento pueden pesarse simultáneamente no más de k monedas. Todas las monedas de X son iguales, a excepción de una moneda que es falsa y difiere de las demás por el peso. Se requiere ofrecer un algoritmo estático de la búsqueda, que en un número mínimo posible de experimentos revele la moneda falsa.

Convengamos en considerar, para mayor precisión, que los pesos de las monedas auténticas son iguales a 0, y el de la moneda falsa, a la unidad. Supongamos que en el i -ésimo experimento ($i = 1, \dots, N$) se determina el peso sumario de los subconjuntos de monedas $Y_i \subseteq X$, $|Y_i| \leq k$. Está claro que el sistema $\{Y_1, \dots, Y_N\}$ de los subconjuntos de X representa la estrategia estática de la búsqueda, si y sólo si este sistema es separador. El problema se ha reducido, pues, a la construcción de un sistema separador mínimo, cada subconjunto del cual contiene no más de k elementos.

Designaremos con un mismo símbolo x_j la j -ésima moneda y el peso de la misma; $j = 1, \dots, n$. Igual que antes, construyamos, a base del juego de subconjuntos $\{Y_1, \dots, Y_N\}$, una $(N \times n)$ -matriz de los experimentos $A = [a_{ij}]$, al hacer $a_{ij} = 1$, si $x_i \in Y_j$, y $a_{ij} = 0$, en el caso contrario. Observemos que cada fila de la matriz A contiene no más de k unidades, con la particularidad de que el sistema $\{Y_1, \dots, Y_N\}$ es separador, si y sólo si todas las columnas de la matriz A son diferentes. Tal $(0, 1)$ -matriz A lleva el nombre de $F(n, \leq k)$ -plan. En este caso, si todas las filas de A contienen exactamente k unidades, la matriz A se denomina $F(n, k)$ -plan. No es difícil ver que las cantidades mínimas de filas en $F(n, k)$ y, correspondientemente, en $F(n, \leq k)$ -plan son iguales a $N(n, k)$ y $N(n, \leq k)$.

Procedamos a estudiar estas magnitudes y construir los planes correspondientes.

Teorema 11. Son válidas las siguientes afirmaciones:

- 1) $N(n, k) = N(n, n - k)$,
- 2) $N(n, k) = N(n, \leq k)$ para $n \geq 2k$.

Demostración. La validez de la primera afirmación se deduce de que todo $F(n, k)$ -plan se transforma, por inversión de 0 y 1, en $F(n, n - k)$ -plan. Demostremos 2). Por cuanto todo $F(n, k)$ -plan es, a la vez, el $F(n, \leq k)$ -plan, entonces $N(n, \leq k) \leq N(n, k)$. La desigualdad inversa para $n \geq 2k$ se deduce del algoritmo que convierte el $F(n, \leq k)$ -plan en $F(n, k)$ -plan. En efecto, veamos una fila arbitraria del $F(n, \leq k)$ -plan y denotemos con m el número de unidades en ella. Consideremos las columnas del plan que

en la fila elegida contienen cero. Está claro que entre estas $n - m$ columnas pueden elegirse no menos de $(n - m) - m = n - 2m$ tales columnas que la sustitución en ellas del elemento nulo destacado por la unidad conserve intacta la propiedad de distinguibilidad de las columnas del plan. Aprovechando esta circunstancia, completemos en cada fila el número de unidades del $F(n, \leq k)$ -plan que faltan para obtener k unidades y, de este modo, transformémoslo en $F(n, k)$ -plan.

Los resultados de los siguientes dos teoremas los obtuvo el matemático húngaro Katona [37].

Teorema 12.

$$\left\lceil \frac{2(n-1)}{k+1} \right\rceil \leq N(n, \leq k) \leq \max \left\{ \left\lceil \frac{2(n-1)}{k+1} \right\rceil, k \right\}.$$

Demostración. Por definición del $F(n, \leq k)$ -plan, la cantidad total de unidades en el plan no sobrepasa de Nk . Sea n_i el número de columnas del plan que contienen i unidades. Por cuanto $n_0 \leq 1$, $n_1 \leq N$, tenemos

$$Nk \geq \sum_{i=0}^N i n_i \geq n_1 + 2 \sum_{i=2}^N n_i = n_1 + 2(n - n_1 - n_0) \geq 2(n-1) - N,$$

de donde se desprende precisamente la estimación inferior. Para demostrar

la estimación superior, pongamos $N = \max \left\{ \left\lceil \frac{2(n-1)}{k+1} \right\rceil, k \right\}$ y

mostremos que existe un $F(n, \leq k)$ -plan de dimensión $N \times n$.

Examinaremos diferentes columnas binarias de altura N . Dos columnas de este género, $\vec{a} = (a_0, \dots, a_{N-1})^T$ y $\vec{b} = (b_0, \dots, b_{N-1})^T$ se llamarán semejantes, si $a_i = b_{(i+j) \bmod N}$ para cierto j natural y para todo $i = 0, \dots, N-1$. El conjunto de todas las columnas se divide en clases de equivalencia con relación a la semejanza. Está claro que la potencia de cada clase de este género será un divisor del número N . Cualesquiera dos columnas semejantes contienen un número igual de unidades. Una matriz compuesta por todas las columnas de una misma clase de equivalencia se denominará bloque.

Veamos un bloque de dimensión $N \times N$, generado por una columna cuyos primeros $m < N$ lugares están ocupados por unidades, y los demás lugares, por ceros. No es difícil notar que para cualquier número natural r , $0 < r < N$, pueden elegirse r columnas de este bloque de un modo tal que el número de unidades en cualesquiera dos filas de la matriz, formada por dichas columnas, se diferenciará no más que en 1. La matriz formada por las citadas r columnas se denominará bloque incompleto.

El $F(n, \leq k)$ -plan se construirá ahora del modo siguiente. Contiene una columna compuesta sólo por ceros y N columnas, cada una de las cuales contiene un solo elemento unidad. Las demás columnas contienen cada una dos elementos unidad. Estos se disponen en el plan en bloques. El

último bloque puede ser incompleto. La cantidad total de columnas que contienen no más de dos unidades no es inferior a n :

$$1 + N + \frac{N(N-1)}{2} = 1 + \frac{N(N+1)}{2} \geq \frac{2(n-1)}{k+1} \cdot \frac{k+1}{2} = n.$$

Por consiguiente, al construir el plan, podemos en realidad servirnos sólo de tales columnas. En el plan construido la cantidad de unidades en cada fila no es superior a k . Efectivamente, la cantidad total de unidades en el plan es igual a $2n - N - 2$. Por cuanto la cantidad de unidades en cualesquiera dos filas se diferencia no más que en 1, cada fila, entonces, contiene no más de $\left\lfloor \frac{2(n-1) - N}{N} \right\rfloor \leq k$ unidades. El teorema está demostrado.

Corolario. $N(n, k) = \left\lfloor \frac{2(n-1)}{k+1} \right\rfloor$ para $n \geq \frac{k(k+1)}{2} + 1$.

Teorema 13. Se verifica la siguiente desigualdad:

$$N(n, k) \geq \frac{\log_2 n}{h(k/n)}$$

donde $h(p) = -p \log_2 p - (1-p) \log_2 (1-p)$.

Demostración. Hagamos uso de la propiedad de subaditividad de la entropía para obtener la estimación inferior de la magnitud $N(n, k)$. Supongamos que la $(N \times n)$ -matriz $A = [a_{ij}]$, donde $N = N(n, k)$, $a_{ij} \in \{0, 1\}$, es el $F(n, k)$ -plan. En este caso A aplica el conjunto B_n^1 de todas las sucesiones de longitud n , que se componen de una sola unidad y ceros, biunivocamente sobre cierto subconjunto del conjunto B_N de todas las $(0, 1)$ -sucesiones de longitud N :

$$\vec{e} = A\vec{x}, \vec{x} = (x_1, \dots, x_n)^T \in B_n^1, \vec{e} = (e_1, \dots, e_N)^T \in B_N.$$

Introduzcamos sobre el conjunto B_n^1 una distribución uniforme $\mathcal{P}$: $P(\vec{x}) = 1/n$, $\vec{x} \in B_n^1$. Debido a ello, sobre el conjunto $A(B_n^1) \subseteq B_N$ se induce una distribución uniforme que se designará con la misma letra $\mathcal{P}$: $P(\vec{e}) = 1/n$, donde $\vec{e} \in A(B_n^1)$. La entropía $H(\mathcal{P})$ de distribución de las probabilidades $\mathcal{P}$ es igual, como es fácil de notar, a $\log_2 n$. Por otra parte, la distribución de las probabilidades $\mathcal{P}$ en $A(B_n^1)$ forma N distribuciones simples de probabilidades $\mathcal{A}_i$, definidas sobre las coordenadas de los vectores de B_N : $P\{e_i = 1\} = k/n$, $P\{e_i = 0\} = 1 - k/n$, $i = 1, \dots, N$. Al calcular las últimas probabilidades se ha utilizado el hecho de que cada fila de la matriz A contiene exactamente k unidades. Al aplicar la propiedad de subaditividad de la entropía, obtenemos: $\log_2 n = H(\mathcal{P}) \leq \sum_{i=1}^N H(\mathcal{A}_i) = Nh(p)$, donde $p = k/n$, $h(p) = -p \log_2 p - (1-p) \log_2 (1-p)$ es la entropía de Shannon. Por consiguiente, $\frac{\log_2 n}{h(p)} \leq N = N(n, k)$.

Corolario. Sea $n \rightarrow \infty$, $k = k(n) = \lceil pn \rceil$, donde el número p no depende de n , y $0 < p < 1/2$. Entonces resulta válida la siguiente fórmula asintótica:

$$N(n, k) = \frac{\log_2 n}{h(p)} (1 + o(1)).$$

Demostración. Construyamos un $F(n, k)$ -plan concreto de longitud $N = \frac{\log_2 n}{h(p)} (1 + o(1))$. Por cuanto cada $F(N, \leq k)$ -plan puede ser transformado, para $n \geq 2k$, en un $F(n, k)$ -plan de la misma longitud será suficiente construir el correspondiente $F(n, \leq k)$ -plan. Elijamos N mínimo natural, para el cual se verifica la desigualdad $\binom{N}{[pN]} \geq n$. Hallando por logaritmos esta desigualdad y empleando la fórmula de Stirling para los factoriales, no es difícil de concluir que para tal N es lícita la fórmula asintótica: $N = \frac{\log_2 n}{h(p)} (1 + o(1))$. El plan se construye a partir de las columnas de altura N , cada una de las cuales contiene exactamente $[pN]$ unidades. Las columnas en el plan se disponen por bloques, empleando en caso de necesidad al final el bloque incompleto. Merced a tal disposición se asegura la diferencia entre el número de unidades en las filas del plan en no más de 1. La desigualdad $\binom{N}{[pN]} \geq n$ garantiza que al construir el plan nos será suficiente la reserva de columnas, con la cantidad de unidades indicada en cada una de ellas. Por cuanto la parte de las unidades en cada columna del plan no es superior a p , mientras que la cantidad de unidades en cualesquiera dos filas se diferencia en no más de 1, cada fila contiene no más de $k = \lceil pn \rceil$ unidades, es decir, dicho plan es realmente $F(n, \geq k)$ -plan. El teorema queda así demostrado.

Aduzcamos una estimación superior para la magnitud $N(n, \leq k)$, la cual es próxima a la estimación inferior obtenida en el teorema anterior por el método de entropía (véase [88 y 89]). Observemos que es un tanto mejor que la estimación superior de Katona de [37].

Teorema 14. Se verifica la desigualdad

$$N(n, \leq k) \leq \left\lceil \frac{\log_2 n}{\log_2 \frac{n}{k}} \right\rceil \left\lceil \frac{n}{k} - 1 \right\rceil.$$

La demostración está basada en la construcción del $N(n, \leq k)$ -plan, partiendo del código binario de Rademajer. Así llamaremos una $(0, 1)$ -matriz de dimensión $\lceil \log_2 n \rceil \times n$, cuyas columnas son todas diferentes. Para obtener algún código de Rademajer es suficiente tomar una potencia del número 2, la más próxima a n superiormente, y, al elegir cualesquiera

n números inferiores a la citada potencia, escribirlos en columnas conforme al sistema de numeración con base 2.

Sea $n = 3$. Construyamos para este caso todos los códigos de Rademajer equivalentes con una exactitud de hasta una permutación de las columnas:

$$\begin{array}{cccccc} 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{array}$$

Nos hará falta la siguiente identidad sencilla, válida para cualquier número real x y natural n :

$$\left\lceil \left\lfloor \frac{x}{n} \right\rfloor \right\rceil = \left\lfloor \frac{x}{n} \right\rfloor. \quad (9)$$

Demostración del teorema. Sean $R = R(k)$ un código arbitrario de Rademajer de longitud k , t , un número natural y $s_0, \dots, s_{t-1}$, números arbitrarios del conjunto $\{0, 1\}$. Partiendo de R , construyamos otro código de Rademajer $R_{s_0}, \dots, R_{s_{t-1}}$ del modo siguiente. Tomemos cada fila con número $i \in \{1, \dots, \lceil \log_2 k \rceil\}$ y sustituylamos en ella cada número $l \in \{0, 1\}$ por el número $(l + s_{(i-1) \bmod t}) \bmod 2$.

Escribamos arbitrariamente uno tras otro todos los 2^t códigos $R_{s_0}, \dots, R_{s_{t-1}}$ y obtendremos una matriz M de dimensión $\lceil \log_2 k \rceil \times 2^t k$. Esta matriz se divide en $\lambda_0 = \lceil t^{-1} \rceil \lceil \log_2 k \rceil = \lceil t^{-1} \log_2 k \rceil$ submatrices M_λ , $\lambda = 1, \dots, \lambda_0$ (véase (9)). Las submatrices M_λ están formadas, para $\lambda \neq \lambda_0$, por t filas de la matriz M con los números $i = t(\lambda - 1) + 1, t(\lambda - 1) + 2, \dots, t\lambda$. La submatriz M_{λ_0} está formada por las filas con los números $i = t(\lambda_0 - 1) + 1, t(\lambda_0 - 1) + 2, \dots, \lceil \log_2 k \rceil$ y puede tener menos de t filas.

No es difícil notar que cualquier submatriz M_λ contiene, para $\lambda \neq \lambda_0$, cada una de las 2^t columnas binarias de altura t exactamente k veces. Designemos el número de filas de la submatriz M_{λ_0} con τ . Análogamente, M_{λ_0} contiene cada una de las 2^t columnas binarias de altura τ exactamente $2^{t-\tau} \cdot k$ veces.

Denotemos con L_μ una matriz compuesta de k columnas iguales de altura t que representa una notación en columna del número $\mu \in \{0, 1, \dots, 2^t - 1\}$ con arreglo al sistema binario de numeración. Escribamos sucesivamente, una tras otra, todas las 2^t matrices L_μ ; designemos con M_0 la matriz obtenida. Agreguemos superiormente a la matriz M la matriz M_0 y designemos con $\tilde{M}$ el resultado.

La matriz $\tilde{M}$ contiene $t + \lceil \log_2 k \rceil$ filas y $2^t \cdot k$ columnas que son distintas dos a dos. Por consiguiente $\tilde{M}$ es el código de Rademajer.

La demostración del teorema se basa en la construcción de un $F(n, \leq k)$ -plan, partiendo del código $\tilde{M}$, donde el número natural t se determina unívocamente de la condición $2^{t-1} < n/k \leq 2^t$. Para cada $\lambda = 0, \dots, \lambda_0 - 1$ construyamos la matriz M'_λ compuesta de n columnas de altura

$\left\lceil \frac{n}{k} \right\rceil - 1$. En cada columna todos los elementos son nulos, a excepción, quizás, de un elemento igual a 1, y en M'_λ cada columna no nula figura no más de k veces. En este caso dos columnas en M'_λ son iguales cuando y sólo cuando lo son las columnas con los números correspondientes en M_λ . Observemos que mediante dichas propiedades la matriz M' se define no unívocamente. Para nosotros tiene importancia el hecho de que tal matriz siempre existe. Está claro que cada fila de M'_λ contiene no más de k unidades.

De un modo análogo construyamos la matriz M'_{λ_0} . Esta consta de n columnas de altura $\left\lceil \frac{n}{k} \right\rceil - 2^{t-1}$. Cada una de estas columnas es o bien íntegramente nula, o bien contiene una sola unidad. Cada columna no nula figura en M'_{λ_0} no más de k veces, mientras que la columna nula, no más de $2^{t-1} \cdot k$ veces. Las columnas en M'_{λ_0} se disponen de un modo tal que de la condición de igualdad de dos columnas en M'_{λ_0} se deduce la igualdad de las columnas correspondientes en M'_{λ_0} . Cada fila de M'_{λ_0} contiene también no más de k unidades.

Escribamos sucesivamente, una debajo de la otra, las matrices $M'_0, M'_1, \dots, M'_{\lambda_0}$ y designemos el resultado con $\tilde{M}'$. Todas las columnas en $\tilde{M}'$, y, por tanto, en M' son distintas de dos en dos. Cada fila de $\tilde{M}'$ contiene k unidades a lo sumo. Por consiguiente, $\tilde{M}'$ es un $F(n, \leq k)$ -plan. Estimemos su longitud N :

$$\begin{aligned} N &= \lambda_0 \left(\left\lceil \frac{n}{k} \right\rceil - 1 \right) + \left(\left\lceil \frac{n}{k} \right\rceil - 2^{t-1} \right) \leq (\lambda_0 + 1) \times \\ &\times \left(\left\lceil \frac{n}{k} \right\rceil - 1 \right) = \left\lceil \frac{n}{k} \right\rceil - 1 + \frac{\log_2 k}{t} - \left\lceil \left(\left\lceil \frac{n}{k} \right\rceil - 1 \right) \right\rceil \leq \\ &\leq \left\lceil \frac{n}{k} \right\rceil + 1 + \frac{\log_2 k}{\log_2 \frac{n}{k}} \left\lceil \frac{n-k}{k} \right\rceil = \left\lceil \frac{\log_2 n}{\log_2 \frac{n}{k}} \right\rceil \left\lceil \frac{n-k}{k} \right\rceil. \end{aligned}$$

El teorema está demostrado.

Ilustremos la construcción del $F(n, \leq k)$ -plan con un ejemplo en que $n = 20$ y $k = 5$. Tenemos: $t = \left\lceil \log_2 \frac{n}{k} \right\rceil = 2$. Veamos el código de Rademajer R de longitud 5:

$$R: \begin{matrix} 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 \end{matrix}$$

Determinemos $R_{00}, R_{01}, R_{10}, R_{11}, L_0, L_1, L_2, L_3$:

$$\begin{array}{lll}
 R_{00}: \begin{array}{l} 1\ 1\ 1\ 0\ 0 \\ 1\ 1\ 0\ 1\ 0; \\ 1\ 0\ 1\ 1\ 0 \end{array} & R_{01}: \begin{array}{l} 1\ 1\ 1\ 0\ 0 \\ 0\ 0\ 1\ 0\ 1; \\ 1\ 0\ 1\ 1\ 0 \end{array} & R_{10}: \begin{array}{l} 0\ 0\ 0\ 1\ 1 \\ 1\ 1\ 0\ 1\ 0; \\ 0\ 1\ 0\ 0\ 1 \end{array} \\
 R_{11}: \begin{array}{l} 0\ 0\ 0\ 1\ 1 \\ 0\ 0\ 1\ 0\ 1; \\ 0\ 1\ 0\ 0\ 1 \end{array} & L_0: \begin{array}{l} 0\ 0\ 0\ 0\ 0 \\ 0\ 0\ 0\ 0\ 0; \end{array} & L_1: \begin{array}{l} 0\ 0\ 0\ 0\ 0; \\ 1\ 1\ 1\ 1\ 1; \end{array} \\
 L_2: \begin{array}{l} 1\ 1\ 1\ 1\ 1 \\ 0\ 0\ 0\ 0\ 0; \end{array} & L_3: \begin{array}{l} 1\ 1\ 1\ 1\ 1 \\ 1\ 1\ 1\ 1\ 1. \end{array}
 \end{array}$$

Las construcciones ulteriores del código $\tilde{M}$ y del $F(20, 5)$ -plan se ven del diagrama siguiente.

Código $\tilde{M}$:

$$\begin{array}{l}
 M_0: \begin{array}{llll} 0\ 0\ 0\ 0\ 0 & 0\ 0\ 0\ 0\ 0 & 1\ 1\ 1\ 1\ 1 & 1\ 1\ 1\ 1\ 1 \\ 0\ 0\ 0\ 0\ 0 & 1\ 1\ 1\ 1\ 1 & 0\ 0\ 0\ 0\ 0 & 1\ 1\ 1\ 1\ 1 \end{array} \\
 M: \left\{ \begin{array}{l} M_1: \begin{bmatrix} 11100 & 11100 & 00011 & 00011 \\ 11010 & 00101 & 11010 & 00101 \end{bmatrix} \\ M_2: \begin{bmatrix} 10110 & 10110 & 01001 & 00101 \end{bmatrix} \end{array} \right.
 \end{array}$$

$\Gamma(20, 5)$ -plan

$$\begin{array}{l}
 M_0': \begin{bmatrix} 1\ 1\ 1\ 1\ 1 & 0\ 0\ 0\ 0\ 0 & 0\ 0\ 0\ 0\ 0 & 0\ 0\ 0\ 0\ 0 \\ 0\ 0\ 0\ 0\ 0 & 1\ 1\ 1\ 1\ 1 & 0\ 0\ 0\ 0\ 0 & 0\ 0\ 0\ 0\ 0 \\ 0\ 0\ 0\ 0\ 0 & 0\ 0\ 0\ 0\ 0 & 1\ 1\ 1\ 1\ 1 & 0\ 0\ 0\ 0\ 0 \end{bmatrix} \\
 M_1': \begin{bmatrix} 1\ 1\ 0\ 0\ 0 & 0\ 0\ 1\ 0\ 0 & 0\ 0\ 0\ 1\ 0 & 0\ 0\ 0\ 0\ 1 \\ 0\ 0\ 1\ 0\ 0 & 1\ 1\ 0\ 0\ 0 & 0\ 0\ 0\ 0\ 1 & 0\ 0\ 0\ 1\ 0 \\ 0\ 0\ 0\ 1\ 0 & 0\ 0\ 0\ 0\ 1 & 1\ 1\ 0\ 0\ 0 & 0\ 0\ 1\ 0\ 0 \end{bmatrix} \\
 M_2' \left\{ \begin{array}{l} 1\ 0\ 1\ 1\ 0 & 1\ 0\ 1\ 0\ 0 & 0\ 0\ 0\ 0\ 0 & 0\ 0\ 0\ 0\ 0 \\ 0\ 0\ 0\ 0\ 0 & 0\ 0\ 0\ 1\ 0 & 0\ 1\ 0\ 0\ 1 & 0\ 1\ 0\ 0\ 1 \end{array} \right.
 \end{array}$$

De los teoremas 13 y 14 se deduce un

Corolario. Sea $n \rightarrow \infty$, $k = k(n) = o(n)$. En este caso es válida la fórmula asintótica:

$$N(n, k) = \frac{n \log_2 n}{k \log_2 \frac{n}{k}} (1 + o(1)).$$